1568

Critical Reviews in Toxicology

http://informahealthcare.com/txc
ISSN: 1040-8444 (print), 1547-6898 (electronic)

Crit Rev Toxicol, 2014; 44(S3): 17–43
 DOI: 10.3109/10408444.2014.931925

informa
healthcare

REVIEW ARTICLE

The use of mode of action information in risk assessment: Quantitative key events/dose-response framework for modeling the dose-response for key events

Ted W. Simon[1], S. Stoney Simons Jr.[2], R. Julian Preston[3], Alan R. Boobis[4], Samuel M. Cohen[5], Nancy G. Doerrer[6], Penelope A. Fenner-Crisp[7], Tami S. McMullin[8], Charlene A. McQueen[9], J. Craig Rowlands[10], and RISK21 Dose-Response Subteam[11]

[1]*Ted Simon LLC, Winston, GA, USA,* [2]*Steroid Hormones Section, NIDDK/LERB, National Institutes of Health, Bethesda, MD, USA,* [3]*U.S. EPA/National Health and Environmental Effects Research Laboratory, Research Triangle Park, NC, USA,* [4]*Imperial College London, London, UK,* [5]*University of Nebraska Medical Center, Omaha, NE, USA,* [6]*ILSI Health and Environmental Sciences Institute, Washington, DC, USA,* [7]*Consultant, North Garden, VA, USA,* [8]*Dow Corning, Midland, MI, USA,* [9]*U.S. EPA/National Health and Environmental Effects Research Laboratory, Research Triangle Park, NC, USA, and* [10]*The Dow Chemical Company, Toxicology and Environmental Research & Consulting, Midland, MI, USA,* [11]*ILSI Health and Environmental Sciences Institute, Washington, DC, USA.*

Abstract

The HESI RISK21 project formed the Dose-Response/Mode-of-Action Subteam to develop strategies for using all available data (*in vitro, in vivo,* and *in silico*) to advance the next-generation of chemical risk assessments. A goal of the Subteam is to enhance the existing Mode of Action/Human Relevance Framework and Key Events/Dose Response Framework (KEDRF) to make the best use of quantitative dose-response and timing information for Key Events (KEs). The resulting Quantitative Key Events/Dose-Response Framework (Q-KEDRF) provides a structured quantitative approach for systematic examination of the dose-response and timing of KEs resulting from a dose of a bioactive agent that causes a potential adverse outcome. Two concepts are described as aids to increasing the understanding of mode of action—Associative Events and Modulating Factors. These concepts are illustrated in two case studies; 1) cholinesterase inhibition by the pesticide chlorpyrifos, which illustrates the necessity of considering quantitative dose-response information when assessing the effect of a Modulating Factor, that is, enzyme polymorphisms in humans, and 2) estrogen-induced uterotrophic responses in rodents, which demonstrate how quantitative dose-response modeling for KE, the understanding of temporal relationships between KEs and a counterfactual examination of hypothesized KEs can determine whether they are Associative Events or true KEs.

Abbreviations: *AChE* acetyl cholinesterase, *AE* associative event, *AOP* adverse outcome pathway, *As3mt* arsenic methyltransferase, *AUC* area under the curve, *BMDL* benchmark dose lower confidence limit, *BMR* benchmark response, *BPA* bisphenol A, *BrdU* bromodeoxyuridine, *BuChE* butyrylcholinesterase, *ChE* cholinesterase, *CPF* chlorpyrifos, *CYP450* cytochrome P450, *DES* diethylstilbestrol, DMA^{III} dimethylarsinic acid (reactive metabolite trivalent), DMA^{V} dimethylarsinic acid, *DMPS* dimercaptopropanesulfonic acid, *DR* dose-response, *EC* European Commission, EC_{50} median effective concentration, *EFSA* European Food Safety Authority, *EPA* Environmental Protection Agency (US), *ER* estrogen receptor, *ERα* estrogen receptor alpha, HESI Health and Environmental Sciences Institute, *HRF* Human Relevance Framework, *ILSI* International Life Sciences Institute, *IVIVE in vitro to in vivo* extrapolation, *KE* key event, *KEDRF Key* Events/Dose-Response Framework, *L-NAME* L-NG-nitroarginine methyl ester, *ModF* modulating factor, *MIE* molecular initiating event, *MOA* mode of action, *MOE* margin of exposure, *NOAEL* no-observed-adverse-effect level, *NRC* National Research Council, *OECD* Organisation for Economic Co-operation and Development, *OP* organophosphate, *PBPK* physiologically based pharmacokinetic, *PD* pharmacodynamic, *PON1* Paraoxonase 1, *PR* progesterone receptors, *Q-KEDRF* Quantitative Key Events/Dose-Response Framework, *QSAR* quantitative structure-activity relationship, *RBC* red blood cell, *REACH* Registration, Evaluation, Authorisation and Restriction of Chemicals, *RIP140* receptor interacting protein 140, *SAM* S-adenosyl methionine, *SRC-1* steroid receptor coactivator-1, *TCPy* 3,5,6-trichloro-2-pyridinol, *TDV* traditional dose value, *WoE* weight of evidence.

Keywords

associative event, key event, mode of action, modulating factor, Q-KEDRF, risk assessment

History

Received 24 January 2014
Revised 22 May 2014
Accepted 3 June 2014
Published online 7 July 2014

Address for correspondence: Nancy G. Doerrer, ILSI Health and Environmental Sciences Institute, 1156 15th Street, NW, Washington, DC 20005, USA. Tel: +1 202-659-3306. E-mail: ndoerrer@hesiglobal.org; URL: http://www.risk21.org

Table of Contents

Introduction

As society progresses through the second decade of the 21st century, there is increased pressure to embrace new ideas and new information in the practice of toxicology and risk assessment. Modern biological science has provided many assessment tools—genomics, transcriptomics, proteomics, metabolomics, and others—that enable scientists to dissect and ultimately understand the biological pathways underlying toxicity. Disruption of these pathways is associated with adverse outcomes.

The progression of this understanding of these adverse outcome pathways fosters and enables the use of these new tools in the practice of chemical risk assessment (Ankley et al. 2010, NRC 2007). What is needed is the knowledge of the biological pathways that underlie a given toxicity and an estimate of the degree or amount of disruption each pathway can tolerate without the occurrence of pathway-specific toxicity (Boekelheide and Andersen 2010, Boekelheide and Campion 2010, Hartung and McBride 2011). The use of mode of action (MOA) currently is the most reliable way for developing sufficient knowledge and understanding of these biological pathways.

RISK21 project

For a number of years, the International Life Sciences Institute (ILSI) Research Foundation has assembled cross-disciplinary working groups to examine current risk assessment approaches for evaluating dose-response and identifying safe exposure levels (Julien et al. 2009). Recently, these efforts were applied to four categories of bioactive agents—food allergens, nutrients, pathogenic microorganisms, and environmental chemicals—and from the lessons learned, a common analytical framework was developed for understanding MOA—the Key Events/Dose-Response Framework (KEDRF; Boobis et al. 2009, Buchanan et al. 2009, Julien et al. 2009, Ross et al. 2009, Taylor et al. 2009).

The present paper describes ways to incorporate information about the timing of occurrence and quantitative dose-response of Key Events (KE) into the KEDRF. This expanded framework is known as the Quantitative Key Events/Dose-Response Framework or Q-KEDRF. In one sense, this is a "how-to" paper, which describes methods to incorporate additional information for understanding the particulars of the MOA of a chemical. In addition to a discussion of these methods, examples are provided for illustration.

Dose-response/Mode-of-Action Subteam

A central issue in 21st century toxicology and risk assessment is dose-response analysis and its extrapolation to human exposure levels. Building on the KEDRF, the Dose-Response (DR)/Mode-of-Action (MOA) Subteam within the ILSI Health and Environmental Sciences Institute's (HESI's) RISK21 project was formed to develop a clear strategy for using all available data (*in vitro, in vivo,* and *in silico*) in both qualitative and quantitative ways to develop the methods to be used in next-generation risk assessments of substances. The gathering of these various types of data is best accomplished in a tiered fashion suggested by the red triangle labeled as "Toxicity" in the upper left portion of Figure 1.

The DR/MOA Subteam has three main objectives: 1) to provide a forum to discuss approaches to dose extrapolation in human health risk assessment; 2) to address how an understanding of MOA will influence low-dose extrapolation; and 3) to enhance the existing MOA/Human Relevance Framework (HRF) and KEDRF. Specifically, this third objective aims to use quantitative dose-response and temporal information about both KEs and the adverse outcome in a more robust way. Consistent with all HESI projects, participation in the Risk21

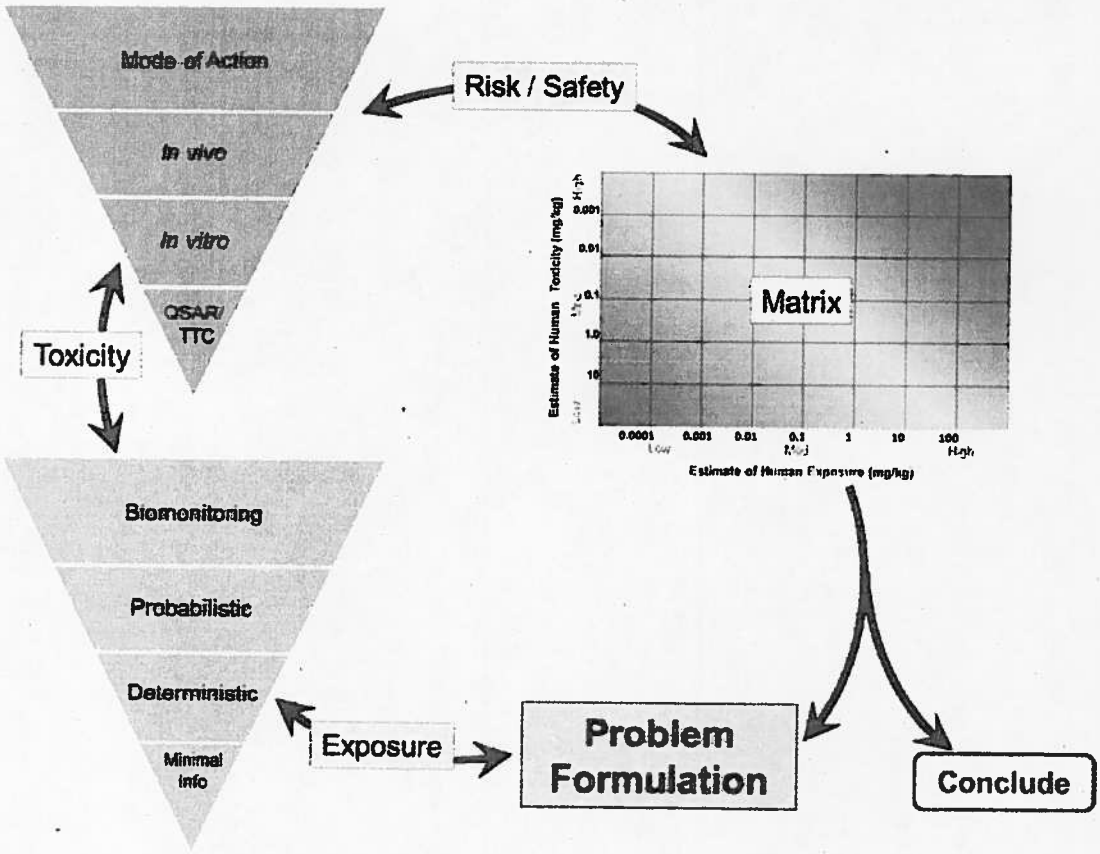


Figure 1. The HESI RISK21 Roadmap and Matrix.

Dose-Response Subteam included tripartite representation from government, academia, and industry, with subteam co-leadership provided by expert scientists from academia and industry.

History and uses of MOA/HRF frameworks

MOA is defined specifically in the US Environmental Protection Agency's (EPA's) 2005 *Guidelines for Carcinogen Risk Assessment* as follows:

> *... a sequence of Key Events and processes, starting with interaction of an agent with a cell, proceeding through operational and anatomical changes, and resulting in cancer formation. A "key event" is an empirically observable precursor step that is itself a necessary element of the mode of action or is a biologically based marker for such an element. Mode of action is contrasted with "mechanism of action," which implies a more detailed understanding and description of events* (USEPA 2005a).

While necessary, single KEs by themselves are not usually sufficient for the adverse outcome to occur, as noted by Julien et al. (2009):

> *Hence, a key event is a necessary, though not a sufficient, step in a process that results in a specific adverse effect.*

Julien et al. (2009) also provide some historical perspective on the concept of MOA and broadened the definition as the "fundamental biological events and processes that underlie the effect of a bioactive agent". In risk assessment, consideration of MOA likely originated from the work of Lehman-McKeeman et al. (1989) on male rat nephrotoxicity associated with accumulation of alpha 2μ-globulin, the work of Cohen and Ellwein (1990) and Cohen (1995) on bladder carcinogenesis, and that of Faustman et al. (1997) on the evaluation of mechanisms of developmental toxicity.

The KEDRF provides a structured approach for systematic examination of KEs that occur between the initial dose of a bioactive agent and the final or apical effect of concern (Julien et al. 2009). Here, not only are the timing of KEs and the quantitative aspects of dose-response examined, but also two additional concepts for understanding MOA are discussed—Associative Events (AEs) and Modulating Factors (ModFs). These concepts were defined in Andersen et al. (2014). AEs essentially provide biomarkers for KEs, and a full definition is provided in a later section. ModFs affect the timing and/or dose-response of KEs and include variability in homeostasis or repair capacities, adaptive or immune mechanisms, enzyme polymorphisms, and other biological factors. The nature and strength of ModFs varies between individuals and in the same individual over time. Life stage, disease state, genetics, lifestyle, and other factors underlie this inter- and intra-individual variability. The Q-KEDRF provides a means to incorporate ModFs in specific situations (described below), and thus, to understand how these result in distributions of population sensitivity in the dose-response of the various KEs and, ultimately, the adverse outcome.

MOA included in regulatory guidance

Government regulatory agencies around the world have incorporated MOA/HRFs into guidance documents because of their ability to inform risk assessments. For example, the European Commission (EC) has incorporated MOA in its risk assessment guidance for industrial chemicals and biocides, and the US EPA's *Guidelines for Carcinogen Risk Assessment* specifically emphasizes the use of MOA information for interpreting and quantifying the potential cancer risks to humans (EC-JRC 2003, USEPA 2005a). In addition, EPA's *Supplemental Guidance for Assessing Susceptibility from Early-Life Exposure to Carcinogens* (or Supplemental Guidance) also relies on knowledge of the MOA (USEPA 2005b). The EPA has also drafted a *Framework for Determining a Mutagenic Mode of Action for Carcinogenicity* that is also based upon MOA, but this guidance has not yet been finalized (USEPA 2007). Health Canada considers MOA in development of drinking water guidelines and pesticide resistance management labeling (Health Canada 1999, 2009, 2011, Liteplo and Meek 2003).

The European Food Safety Authority (EFSA) includes a MOA assessment in its guidance on Harmonizing Cancer and Non-cancer Risk Assessment Approaches (EFSA 2005). MOA is recommended in the EC Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) Regulation guidance for conducting a chemical safety assessment, and in the new "classification, labelling, and packaging" regulation on chemical substances and mixtures (EC 2008). The Organisation for Economic Co-operation and Development (OECD) recommends using MOA to support the building of chemical categories or when using read-across approaches (http://www.oecd.org/chemicalsafety/risk-assessment/groupingofchemicalschemicalcategoriesandread-across.htm). OECD has further embraced the concept of MOA in its recent use of adverse outcome pathways (AOPs; Ankley et al. 2010, OECD 2013). With the push to use more systematic and weight-of-evidence (WoE) approaches in risk assessment, both the recognition of the value and importance of the MOA/HRF and KEDRF and their use in risk assessments will increase.

MOA reduces uncertainty and informs quantitative risk assessment

MOA is a fundamental component of risk assessment for the classification of carcinogens and systemic toxicants, and informing the choice of whether a nonlinear or linear approach to low-dose extrapolation is appropriate. Evaluators can use quantitative kinetic and/or dynamic data considered in MOA analysis in at least five ways. These are listed below, along with specific examples:

1) replace default species extrapolation factors;
2) evaluate more directly the relevant concentrations in the target tissue;
3) determine the most representative dose metric;
4) choose the most appropriate quantitative dose-response model; and
5) assess quantitatively the overall relevance to humans.

Replacement of the default toxicodynamic component of the species extrapolation factor was based on species-dependent differences in the dose-response for AHR activation between humans and rodents in a risk assessment for dioxin based on the 2006 NTP cancer bioassay (Budinsky et al. 2014, NTP 2006, Simon et al. 2009). The understanding gained by investigation into the MOA of small intestinal carcinogenesis by hexavalent chromium led to the identification of the flux of hexavalent chromium entering each segment of the small intestine as the best measure of concentration affecting the target tissue (Kirman et al. 2012, Thompson et al. 2014). The extensive work on the MOA of the pesticide chlorpyrifos (discussed in detail below) enabled the recent identification of brain cholinesterase inhibition as the most appropriate dose metric for a risk assessment based on cholinesterase inhibition (Reiss et al. 2012). An examination of the MOA of acrylamide-induced mammary tumors in F344 rats suggested that nonlinear low-dose extrapolation was a more appropriate method than linear extrapolation (Maier et al. 2012). Last, the Q-KEDRF is part of the MOA/human relevance framework (MOA/HRF) and the purpose of this larger framework is the assessment of human relevance (Boobis et al. 2006, Boobis et al. 2008, Cohen et al. 2003, Cohen et al. 2004, Cohen and Arnold 2011, Meek et al. 2003, Meek 2008, Seed et al. 2005, Meek et al. 2014a, Meek et al. 2014b).

An understanding of MOA is also needed to account for the role of metabolism in various tissues and to decide which early metabolic changes may be KEs. This understanding enables the evaluator to account for induction or inhibition of metabolism of a particular chemical and for potential first-pass effects that may increase or decrease toxicity due to metabolite formation or reduction in the systemic dose of the parent compound. Variations in patterns of toxicity with different metabolic profiles exist across species, strains and sexes in animals and across potentially susceptible subgroups and different life stages in humans. These variations need to be considered so that appropriate and defensible quantitative adjustments can be made for purposes of incorporation of these differences into risk assessments. The overall result is that MOA information can reduce uncertainties in risk assessments in a number of areas.

MOA is the foundation of 21st century toxicology testing and risk assessment

The interpretation of traditional animal toxicity studies for their relevance to humans is difficult, at times impossible, and, more often than not, fraught with controversy (Seok et al. 2013, Beyer et al. 2011, Gori 2013, NRC 1983). These studies generally use high doses resulting in considerable uncertainty when attempting to extrapolate the effects observed in animals to humans, especially when humans are experiencing much lower environmental exposures (NRC 1983). Aspects of this interpretation no less important than human relevance include: 1) the advances in understanding MOA, including the molecular and cellular events responsible for toxicity; 2) the desire to refine, reduce and replace the use of animals in regulatory toxicity testing; and 3) the need for toxicity evaluations for the large number of chemicals in commercial use. In response to these issues, the National Research Council (NRC) developed recommendations on toxicity testing that incorporated new *in vitro* and *in silico* technologies and computational systems biology to complement, and eventually replace, whole animal testing. The new strategy was presented in a report titled

Toxicity Testing in the 21st Century: a Vision and a Strategy (NRC 2007).

The report emphasized the importance of relating events leading to toxicity in the context of perturbations in biologic functions, some of which may be reversible or may represent biologically appropriate adaptations to stressors. Twenty–first century risk assessment uses the knowledge of MOA to link together perturbations in biological pathways observed in humans, in animals, in experiments with *in vitro* systems, and even those predicted by quantitative structure–activity relationships (QSAR) or other computational methods with the goal of determining the likelihood of adverse health outcomes in humans (upper left box in Figure 2).

One vital aspect of this new strategy and the vision of 21st century risk assessment is the development of appropriate prediction models (Adeleye et al. 2014, Judson et al. 2014, Patlewicz et al. 2013). Statistical approaches that attempt to correlate high throughput assay results with adverse outcomes appear to possess a level of predictivity no better than that derived from chemical structure (Thomas et al. 2012). The realization of this difficulty has fostered the curation of AOPs for use in prediction models (Landesmann et al. 2013, OECD 2013, Vinken 2013). In addition, attempts are being made to develop broad categories of MOAs for the purpose of exploiting extant knowledge across categories in a new application of read across (Briggs et al. 2012, Thomas et al. 2013, Vink et al. 2010). Understanding MOA seems to be a necessary part of eventual use of AOPs for risk assessment. Both dose and time contribute to the development of a biologically adverse response—hence, knowledge of MOA requires a detailed understanding of the dose- and time-dependency of the steps that lead from the initial interaction with a chemical to a specific toxic effect (Rowlands et al. 2014).

The Q-KEDRF—a tool for understanding MOA

MOA provides a link between exposure and the risk of adverse health outcomes—but only when the observed pathway perturbations can be characterized in terms of KEs. An important aspect of the definition of a KE is that its occurrence is necessary for the apical event. The other part of the definition is that a KE is "empirically observable." Necessity, as part of the definition, allows one to develop a counterfactual experiment for a putative KE (Figure 2, Box B2) and actually pose the question of whether it truly is a KE—if the event does not occur, will the adverse outcome occur?

Organizing questions and a toolbox for the Q-KEDRF

Box 1 provides a set of organizing questions for MOA as a prelude to applying the Q-KEDRF for specific MOA analyses. These general questions were developed from the charge questions provided to three expert panels in a workshop held at NIEHS to evaluate nuclear receptor-mediated MOAs for liver carcinogenicity (Budinsky et al. 2014, Corton et al. 2014, Andersen et al. 2014, Elcombe et al. 2014). The questions are sorted into three general areas, but in practice, there will likely be considerable overlap between the questions. Attempting to answer these questions will provide anyone engaged in MOA analysis with an understanding of the extent of knowledge.

Box 2 provides three overall categories of schemes for concise organization of the MOA information resulting from tackling the questions in Box 1. Examples of these methods are given from the papers resulting from the nuclear receptor workshop (Budinsky et al. 2014, Corton et al. 2014, Andersen et al. 2014, Elcombe et al. 2014). Necessarily, the graphical techniques, save for the flow chart, will be quantitative. Although not mentioned specifically

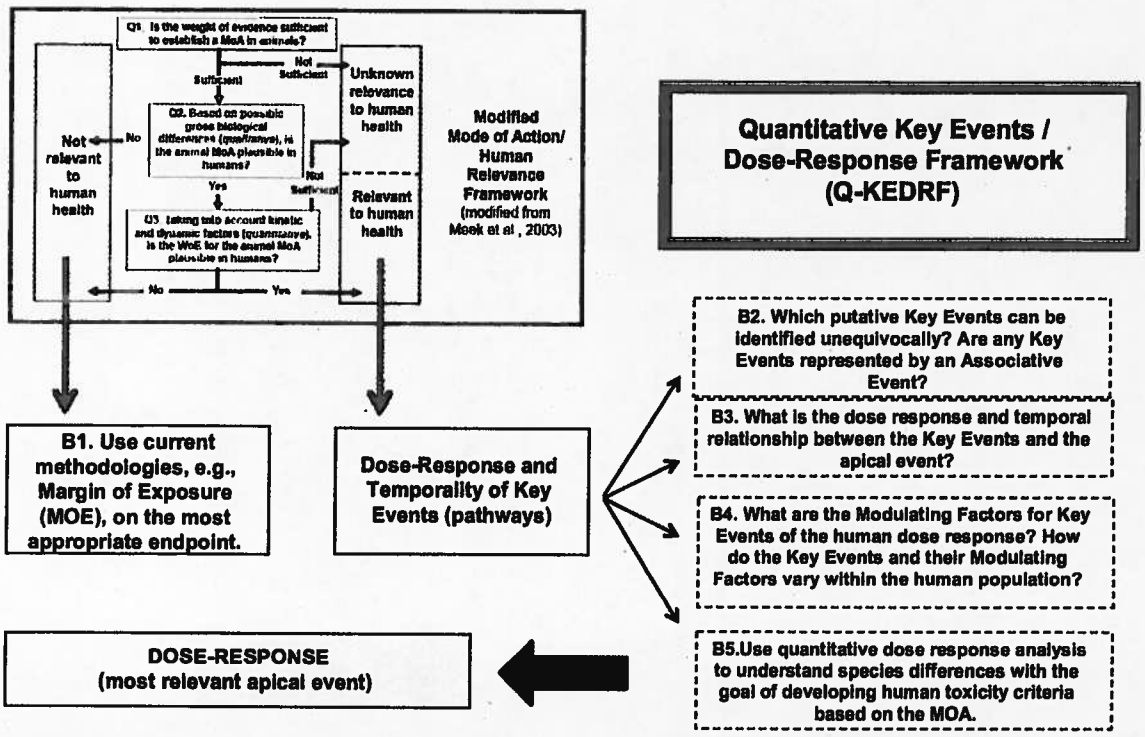


Figure 2. Quantitative Key Events/Dose-Response Framework (Q-KEDRF) and Its Relationship to the Mode of Action/Human Relevance Framework.

Box 1. Organizing questions for mode of action analysis.

Organizing Questions for MOA Consideration

- What is the proposed MOA to be evaluated by the IPCS Human Relevance Framework and modified Bradford Hill considerations?
- Which events are necessary and thus truly key events (KEs)?
- Which events are associative events (AEs)?
- What are the modulating factors (ModFs)?
- Is the proposed MOA likely to be relevant to humans?

Organizing Questions for Quantitative Dose-Response Consideration

- Are extant data sufficient for establishing dose-response relationships for proposed KEs?
- Are extant data sufficient for dose-response modeling of proposed KEs?
- What are the data gaps?
- Does the current understanding support a threshold or non-threshold DR and low dose extrapolation approach?
- On either theoretical or practical grounds, is there a dose or area-under-the-curve (AUC) level insufficient for one or more KEs or the adverse outcome (AO) to occur?

Organizing Questions for Using MOA in Risk Assessment

- Does the weight-of-evidence suggest an appropriate model or approach for the dose-response assessment?
- If so, what are the key data gaps?
- Using a value-of-information (VOI) approach (NRC, 2009; Meek et al., 2014a, b), what data would have the highest value?

in Box 2, exposure–response arrays used in the Toxicological Profiles from the Agency for Toxic Substances and Disease Registry (ATSDR) and the newer Toxicological Reviews from EPA's IRIS program could be organized around proposed KEs within one or more hypothesized MOAs.

WoE considerations for identifying key events and understanding their role in the MOA

Here, we build on the work of Julien et al. (2009) and Andersen et al. (2014) to develop the Q-KEDRF. The following definitions are used in the Q-KEDRF:

Box 2. Overview of the Q-KEDRF toolbox.

Tabular Methods

- Application Scheme for IPCS Human Relevance Framework (Figure 1 in Andersen et al., 2014)
- Comparison of Proposed MOAs (Table 4 in Corton et al., (2014))
- Qualitative Species Concordance Table (Table 4 in Elcombe et al., (2014)
- Qualitative MOA Concordance across Chemicals (Table 5 in Corton et al. (2014)

Graphical Methods

- Flow chart of each proposed MOA (Figure 7; Figure 2 in Corton et al. 2014; Figure 2 in Budinsky et al., 2014)
- Dose-Response Arrays (Figure 8)
- Quantitative Species Concordance Table (Table 3; Table 5 in Budinsky et al., 2014)
- 3D Plotting for Visualizing KEs in Dose and Time (Figure 8 in Budinsky et al., 2014; Figure 6 in Corton et al., 2014)

Quantitative/Computational Methods

- Dose-Response Modeling (BMDS, Graphpad Prism, Other tools)
- Use of Dose Surrogates (AUC, Enzyme Induction levels, etc.)
- Dose-Response Slope Analysis (Tables 6 and 7 here)

DOI 10.3109/10408444.2014.931925

- Key Event (KE): An empirically observable causal precursor step to the adverse outcome that is itself a necessary element of the MOA. KEs are necessary but usually not sufficient for the adverse outcome in the absence of other KEs.
- Associative Events (AEs): Biological processes that by themselves are not KEs in the hypothesized MOA but may serve as reliable indicators or biomarkers for KEs. AEs can be used as surrogates or biomarkers for a KE in a MOA evaluation; depending upon the nature of the biomarker, AEs may reflect exposure to a xenobiotic, the resulting effect, or both.
- Modulating Factors (ModFs): Biological and individual factors, including control mechanisms or host factors, that can modulate the dose–response relationship of one or more KEs, thus altering the probability or magnitude of the adverse outcome (Figure 2, Box B4).

AEs can easily be thought of as biomarkers. In this regard, their relationship to KEs may need to be explored, especially if the AE is needed to measure the KE (IOM 2010).

ModFs may alter the dose-response of the KE in a variety of ways. A selection (not inclusive) of ModFs in humans is provided in Table 1.

Both the KEDRF and Q-KEDRF represent an evolution of the MOA/HRF. Thus, both frameworks assume that sufficient evidence exists to posit the MOA under consideration and to identify hypothesized KEs based on this evidence (Boobis et al. 2006, 2008, 2009, Meek 2008, Meek et al. 2003, Seed et al. 2005, Sonich-Mullin et al. 2001).

If a putative MOA cannot be established, then the Q-KEDRF will not be applicable. Nonetheless, a risk assessment, albeit bearing greater uncertainty, can still be attempted using other methods such as margin of exposure evaluation based on the most appropriate endpoint (Figures 1 and 2, Box B1).

Table 1. Modulating Factors (ModFs) potentially affecting KEs for dose-response in humans. ModFs fall into three general categories shown in the left column. The middle column shows subcategories and the right hand column shows some aspects to consider.

Category	Sub-category	Aspects
Host Factors	Genetic Variation	Polymorphisms
	Disease/Illness	Chronic
		Acute
	Defense mechanisms	Immune responsiveness
		DNA repair
		Cell proliferation
		Cell death
	Physiology	Sex
		Life stage
		ADME
		Hormonal status
Life Style	Diet	Calories
		Fat content
	Tobacco	Usage
	Alcohol	Usage
	Exercise	Frequency
		Intensity
	Pharmaceuticals	Usage
	Illegal drugs	Usage
	Dietary supplements	Vitamins
		Anti-oxidants
Environment	Co-Exposures	Duration
		Air
		Water
		Food
		Dust
		Occupational

A sequence of KEs represents a progression over both dose and time. Knowing the relationship between the various KEs in both dose and time along with an understanding of the underlying biology will contribute to the understanding of the role of particular KE within the MOA. Often, counterfactual information is not available. It may be very difficult to demonstrate the necessity of a particular proposed KE. Understanding the biology can help, but conclusive support of necessity will be a data gap.

Identifying a KE is based on the confidence one has that this event is necessary for the apical event/adverse outcome and is based on an overall WoE evaluation of qualitative and quantitative aspects of the MOA as well as whether the hypothesized roles of the KEs are consistent with the biological basis of the adverse outcome.

The Hill considerations have been adapted for use in understanding MOA. Hill (1965) termed these "viewpoints" or "features to consider" rather than true criteria. Hill's considerations are emphatically not a checklist and necessitate rigorous scientific thinking. They have been quite correctly called "guideposts on the road to common sense" (Phillips and Goodman 2006). Hence, the Key Event/Dose-Response Concordance analysis or Dose-time Concordance analysis requires a rigorous and reasoned WoE approach to reach an understanding of the overall MOA (Phillips and Goodman 2004). Very recently, newly evolved rank-ordered Bradford Hill considerations for application in a MOA analysis were developed (Meek et al. 2014a). In rank order, these include biological concordance, essentiality of key events, concordance of empirical observations, consistency and analogy.

For each proposed KE, if removal or blockade of its occurrence could be accomplished (*i.e.*, the counterfactual experiment), then its necessity (or lack thereof) and consequent identity as a KE could be supported. This is the consideration of essentiality. A cause–effect relationship between a chemical and an adverse effect can never be unequivocally proven because causality itself cannot be proven—only inferred with varying degrees of certainty (Adami et al. 2011). A proposed MOA represents a testable hypothesis (Popper 1959) and the KEs as aspects of that testable hypothesis can be examined in a weight of evidence framework to infer causality (Guzelian et al. 2005, Hill 1965, Phillips and Goodman 2004, 2006, Susser 1986).

Therefore, as indicated in earlier publications on MOA, an essential aspect of the process is identification and evaluation of attendant uncertainties. Each step in a MOA analysis should be accompanied by a list of critical and associated data gaps, with a clear indication of those, if filled, likely to have the most impact on the conclusions. The implications of the existing uncertainties should be explored during dose-response assessment.

Relationships between key events, AEs, and the adverse outcome

The development of a proposed or hypothesized MOA will necessitate identification of KEs and understanding of the dose-response and temporal relationships between the various KEs and the adverse outcome as well as between the KEs

Table 2. Dose-time concordance table for dimethylarsinic acid.

Table —Dose-Time Concordance					
Time	2 weeks	2–3 weeks	10 weeks	25 weeks	104 weeks
Dose (ppm in diet)	Increasing time →				
2	Metabolism*	Metabolism*	Metabolism* Cytotoxicity	Metabolism* Cytotoxicity*	Metabolism* Cytotoxicity*
10	Metabolism*	Metabolism* Cytotoxicity	Metabolism* Cytotoxicity	Metabolism* Cytotoxicity*	Metabolism* Cytotoxicity*
40	Metabolism*	Metabolism* Cytotoxicity	Metabolism* Cytotoxicity Proliferation Hyperplasia	Metabolism* Cytotoxicity* Proliferation* Hyperplasia	Metabolism* Cytotoxicity* Proliferation* Hyperplasia **Carcinomas**
100	Metabolism*	Metabolism Cytotoxicity Proliferation Hyperplasia	Metabolism Cytotoxicity Proliferation Hyperplasia	Metabolism Cytotoxicity* Proliferation Hyperplasia	Metabolism* Cytotoxicity* Proliferation* Hyperplasia **Carcinomas**

The asterisk means that the key event has not been observed at the specific dose/time point but is presumed to have occurred. Although not used here, shading of the table may be helpful with a shading scheme based on the number of KEs. Figure 5 in Meek et al. (2014b) provides another organizational scheme for the dose-time concordance table (Please see Figure 3 for the MOA and text for details).

themselves. This is the purpose of the Dose-Time Concordance table (Table 2). Such a table also addresses the temporal aspects of Box B3 in Figure 2 (Meek et al. 2014b) .

In 2005, EPA's Office of Pesticide Programs proposed a MOA for the carcinogenesis of dimethylarsinic acid or DMA^V, also known as cacodylic acid (USEPA 2005c). DMA^V administered in the diet or drinking water produced bladder cancer in rats. There are four KEs in the MOA for bladder tumors in rats; these are: (1) generation of the reactive metabolite trivalent DMA (DMA^{III}) that is dependent on DMA^V and can be observed as the urinary excretion of trivalent DMA greater than 0.1 µM in urine; (2) cytotoxicity occurring within the superficial epithelial layer of the urinary bladder; (3) consequent regenerative proliferation; and, (4) hyperplasia of the urothelium (Cohen et al. 2006, USEPA 2005c). The qualitative relationships between these KEs in both dose and time is shown in Table 2, which is an example of the dose–time concordance table (Meek et al. 2014a, Meek et al. 2014b).

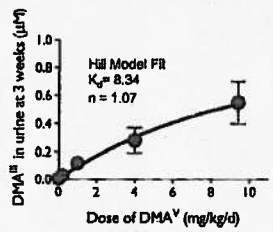
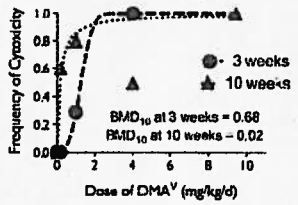
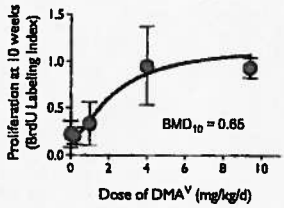
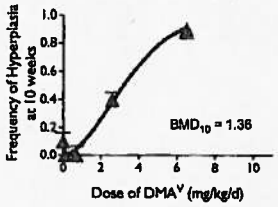
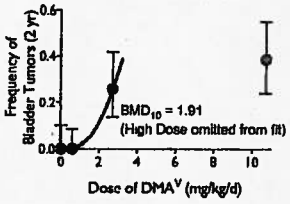
In two-year bioassays, dietary administration of 9.4 mg/kg/d DMA^V produced a statistically significant incidence of tumors; dietary administration of 4.0 mg/kg/d produced a statistically significant incidence of hyperplasia. There were no histopathological changes in the urothelium observable using light microscopy from dietary administration of 1 mg/kg/d or lower. In shorter term mechanistic studies using light and scanning electron microscopies to detect superficial cytotoxic changes, evidence of cytotoxicity was present at dietary doses of 1 mg/kg/d and higher. These same mechanistic studies used bromodeoxyuridine (BrdU) labeling index to assess cell proliferation and observed an increase in proliferation at a dietary dose of 1 mg/kg/d and above.

In rats administered DMA^V in drinking water, genomic microarray analysis revealed a change in the pattern of altered gene expression between 0.4 and 4.7 mg/kg/d, the same level at which an apparent threshold was observed using transmission electron microscopy (Sen et al. 2005). Critical cytotoxic urinary levels of the reactive metabolite DMA^{III} were present in rats orally administered DMA^V at doses of 1 mg/kg/d and above, but absent at 0.2 mg/kg/d. The level of detection for DMA^{III} in urine was 0.01 µM (USEPA, 2005c).

Evidence strongly suggests that DMA^{III} is not DNA reactive, and likely is not genotoxic except at relatively high concentrations (Cohen et al. 2006). Table 2 summarizes the dose-response and temporal relationships for each of the KEs. For risk assessment purposes, it is reasonable to base the assessment on the most sensitive of the KE changes, that is, cytotoxicity. Based on such an analysis, the no-observed-adverse-effect level (NOAEL) is 0.2 mg/kg/d via diet. Similar findings have been identified in rats administered DMA^V in the drinking water (Cohen et al. 2006). Table 3 shows an example of the Dose-Response Species Concordance table that supports quantitative interspecies extrapolation of KEs.

Although the dose-response for humans in Table 3 is lacking, toxicokinetic interspecies extrapolation could be based on differences in the metabolism and kinetics of DMA^V in rats and humans. The evidence indicates that DMA^V is a poor substrate for the methylating enzyme for arsenicals in humans (As^{3+} methyltransferase, As3mt) whereas in rats, this enzyme can readily methylate DMA^V to trimethyl arsenic oxide (Thomas 2007). A physiologically based pharmacokinetic (PBPK) model for DMA^V could support further refinement of the risk assessment, but such a model was not fully developed in 2005 (Evans et al. 2008, USEPA, 2005c). *In vitro* cytotoxicity assays utilizing rat urothelial cells showed an effect at concentrations of approximately 0.2 µM or higher; in comparison, *in vitro* human urothelial cells showed less sensitivity, with cytotoxicity produced at concentrations of 0.5 µM and higher (Cohen et al. 2006). Hence, overall, humans would be less susceptible than rats based on both kinetics and dynamics. These quantitative differences could potentially be used to develop a data-derived species extrapolation factor or chemical-specific adjustment factor (USEPA 2011, WHO-IPCS 2005, Meek et al. 2014b).

Table 3. Dose-Response Species Concordance Table for Key Events (KEs) in the MOA of dimethylarsinic acid (DMA^V) (Adapted from USEPA, 2005c).

	Qualitative concordance			Quantitative concordance and quantitative Dose-response		
Event or factor	Animals	Humans	Concord-ance	Str.*	Animals	Humans
Key events						
Key Event #1 Metabolism to DMA^{III}	DMA^{III} detected in urine following 26 weeks treatment with 100 ppm DMA^V	Evidence following DMA^V exposure too limited to draw conclusions, but DMA^{III} shown to be present following human exposure to iAs	Plausible	+/–		NA
Key Event #2 Urothelial Cytotoxicity	Urothelial toxicity observed in vivo in rats at 2 ppm but not enough for successive key events	Potential to occur in humans but unknown if sufficient DMA^{III} formed	Plausible	+/–		NA
Key Event #3 Urothelial Proliferation	observed at 0.5 mg/kg/d DMA^V	Potential to occur in humans but unknown if sufficient DMA^{III} formed	Plausible	+/–		NA
Key Event #4 Hyperplasia	observed at 2 mg/kg/d or 0.3 to 2 µmol DMA^{III} in urine	Potential to occur in humans but unknown if sufficient DMA^{III} formed	Plausible	+/–		NA
Apical Event Tumors	observed at 5 mg/kg/d DMA^V or 0.8 to 5.05 µmol DMA^{III} in urine	No data in humans	Concordance cannot be made because there is no human data	–		NA

*Str. = strength.

In such a case, this information could be added to the Dose-Response Species Concordance Table.

Low protein or vegetarian diets decrease the availability of S-adenosyl methionine (SAM), and arsenic methylation uses SAM as a methyl donor. Hence, diet may constitute a ModF to be considered (Gamble and Hall 2012).

The risk assessment conducted by EPA's Office of Pesticide Programs (OPP) used a benchmark dose lower confidence limit of 0.07 mg/kg/d DMA^V based on cell proliferation as the 1% point of departure (USEPA 2005c) and a nonlinear low-dose extrapolation to develop a reference dose protective of cancer based on this MOA. Here, this example serves to demonstrate the use of the Dose-Time Concordance Table (Table 2) and the Dose-Response Species Concordance Table (Table 3). The BMD information for KEs occurring at 10 weeks—cytotoxicity, proliferation, and hyperplasia—provided a way to order these KEs and supports their order in the dose–time concordance table (Table 2).

An example of how to use the RISK21 exposure-toxicity matrix is provided (Figure 3). The heavy dotted line on the matrix represents a hazard quotient (HQ) of one. The blue square represents the intersection of exposure and toxicity. If any part of this area extends above the line representing an HQ of one, then exposures may be of concern. In the case of cacodylic acid, all exposure levels within the range of chronic dietary exposures are less than the RfD (USEPA 2006). The exposure-toxicity matrix is flexible; in addition to the range shown here, probability distributions of exposure and/or toxicity can be shown as a means of visualizing probabilistic characterizations of exposure, toxicity, and risk.

Concordance of the MOA between humans and animals

The human relevance of a hypothesized MOA may depend on both qualitative and quantitative factors. As evident from the example with DMA^V above, EPA's Office of Pesticide Pro-

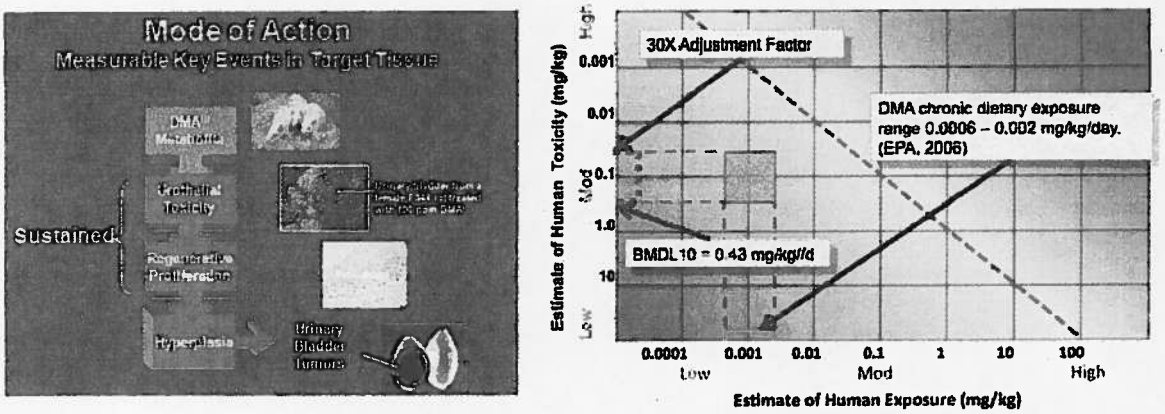


Figure 3. Use of MOA in the HESI RISK21 Matrix. Left: MOA for Tumor Induction by Dimethylarsinic Acid (DMA^V; Cacodylic Acid) that includes cytotoxicity, regenerative proliferation, and hyperplasia. This MOA is used to illustrate the dose-time concordance table and dose-response species concordance table (Tables 2 and 3). Right: Matrix showing the exposure estimates and toxicity range ($BMDL_{10}$ to RfD) for chronic dietary exposure, data from EPA, 2006.

grams clearly recognizes this fact and the need for assessing both qualitative and quantitative concordance of KEs between animals and humans (Dellarco and Fenner-Crisp 2012). For example, in the early 1990s, a technical panel from EPA concluded that male rat renal tubule tumors from chemicals that induced accumulation of α_{2u} globulin were likely not relevant to humans based on qualitative considerations (Rodgers and Baetcke 1993). Naphthalene produces respiratory tract tumors in rats, but the MOA for these tumors in rats is based on metabolic enzyme activity that is not present in humans (Piccirillo et al. 2012).

The Dose-Response Species Concordance Table (Table 3) is a means of illustrating the similarities and differences in a proposed MOA between humans and the test-species. Likely other information, narrative and/or additional tables, will be needed to provide all the information needed for species extrapolation.

Qualitative concordance of key events between humans and animals

Human relevance of the apical endpoint is best determined using a hypothetico-deductive WoE approach (Boobis et al. 2006, 2008, Meek et al. 2003, Rhomberg et al. 2010, Seed et al. 2005, Sonich-Mullin et al. 2001). To address human relevance of the MOA, qualitative concordance between humans and animals for each KE needs to be considered. *In vitro* data from human or animal cells or tissues and/or *in silico* data may also be available; these data play a useful role in the determination of concordance as well. Ideally, the data will be sufficient to determine which of the KEs is relevant to humans, and these data may thus be used to support statements about the relevance to humans of the hypothesized MOA in animals.

Quantitative concordance of the MOA between humans and animals

Quantitative examination of both the dose-response and timing of KEs is also necessary to determine human relevance. For example, a MOA may be operative in both animals and humans, but extremely unlikely in humans because of quantitative toxicokinetic or toxicodynamic differences. If the KE has the potential to occur in humans, then this quantitative examination can be used to inform animal-to-human extrapolation. Hence, the quantitative concordance should provide information about the EC_{50} and/or point-of-departure values for as many KEs as possible in both humans and the animal test species. Including NOAELs or other measures of the no-effect level/threshold such as that defined using the EC_{05} baseline projection method of Silkworth et al. (2005) or the "hockeystick" fitting method of Lutz and Lutz (2009) may also be useful.

The role of quantitative dose-response information

For dose-response assessment, it can be extremely useful to examine quantitative dose-response information from as many relevant sources as possible (*e.g.*, human, laboratory animal or *in vitro* data). These data will help inform the progression of events within the MOA. *In vitro* to *in vivo* extrapolation (IVIVE) may be necessary to express the dose-response for *in vitro* data on a similar dose scale as the *in vivo* data. Where possible, the actual dose-response plots should be shown. It is often helpful to show the dose-response of a KE and that of the apical event or adverse outcome on the same plot (*e.g.*, Figure 2 in Simon, et al 2009). Once the MOA for rat liver tumor promotion by TCDD was considered, the task of arranging the dose-response plots in a figure that displayed the MOA in a meaningful way became easy. Rodent liver tumor promotion is one of the longest and most intensively investigated MOAs in toxicology (Budinsky et al. 2014). Developing similarly informative figures may not be as easy for less well-studied chemicals. Figure 8 is an attempt to create a similar figure for the uterotrophic response. For clarity, it is helpful to have the same dose range on the x-axis in all the plots. When not possible to provide plots of dose-response curves, sufficient narrative should be presented to explain animal/human similarities and differences. If sufficient data in both dose and time are available for a particular KE, a three-dimensional graph with an interpolated surface plot that shows the occurrence of the KE along both dose- and time-axes may be very informative (Box 2; Budinsky et al. 2014).

Use of dose-time and dose-response concordance information in understanding the MOA

In general, events that occur at low doses and/or at early stages in the progression toward the apical event may represent:

- the start of a temporal progression;
- the initial stages of a developing change; or,
- a factor that potentially causes other KEs that occur at higher doses or at a later time in the progression.

Generally, demonstrating that a particular event is necessary is experimentally difficult; yet, it may be possible in some cases (*e.g.*, with transgenic or knockout animals), thus providing a powerful counterfactual demonstration supporting the identification of the event as a KE (Phillips and Goodman 2006). In the example used in Table 2 and Figure 3, let us assume that blocking metabolism of DMA^V or cacodylic acid to dimethyl arsinous acid (DMA^{III}) could reduce or alleviate the KE of urothelial cytotoxicity. The enzyme arsenic methyltranferase (As3mt) catalyzes all steps in the metabolic pathway from arsenite to mono, di, and trimethylated arsenic compounds (USEPA 2005c). If cytotoxicity and tumors did not occur when As3mt was inactivated, this would confirm the role of metabolism and resulting cytotoxicity as necessary and thus as KEs; conversely, if cytotoxicity and tumors occurred even when As3mt was inactivated, one could no longer support the identification of metabolism and cytotoxicity as KEs. Once the DMA^{III} is formed, it readily reacts with free sulfhydryl groups. Co-administration with high doses of a sulfhydryl-containing chemical, such as dimercaptopropanesulfonic acid (DMPS) can act as a trap for the DMA^{III}, reduce or prevent its reaction with proteins, and thus reduce or prevent its biological effects. Co-administration of DMA^V with DMPS inhibits the induction of cytotoxicity and regenerative proliferation of the urinary bladder, providing evidence for DMA^{III} as the reactive intermediate and AE/KE in the DMA^V-induced bladder cancer in rats (Cohen et al. 2006).

The exact nature of a KE cannot be necessarily understood from either its dose-response or its timing of occurrence. For example, some early KEs may need to be sustained in order for later KEs or the apical event/adverse outcome to occur (*e.g.*, Budinsky et al. 2014).

Toxicokinetics may affect this timing. For example, lipid soluble chemicals may be stored in adipose tissue for months or years and produce effects on an ongoing basis; for similar reasons, the dose of a bioaccumulative chemical may be measured as body burden or tissue concentration. In such a case, the area under the curve (AUC) in units of concentration × time would likely represent the ongoing accumulation in both dose and time better than body burden or tissue concentration at a single time point. Sequestration of a chemical by protein binding may also be represented best by the AUC. A monotonic dose–response relationship between the AUC and a biomarker for a putative KE such as enzyme induction indicates that exploring the quantitative relationship between this biomarker and the apical event/adverse outcome may likely help elucidate details of the MOA.

In other cases, the occurrence of some early KEs may trigger a cascade of other events. These early KEs either resolve themselves or are no longer empirically observable. However, the cascade of triggered events continues and leads ultimately to the adverse outcome/apical event. An example of such an effect is illustrated by the difference between long-acting and short-acting estrogens; short-acting estrogens produce early but not late events in the uterotrophic response whereas long-acting estrogens produce both. Estradiol, a long-acting estrogen, can stimulate uterine growth for up to 72 hours whereas the effects of estriol, a short-acting estrogen, last only 24 hours. In fact, estriol and other short-acting estrogens may display partial antagonism when continuously administered in longer-term assays (Clark and Markaverich 1984). Again, these various estrogenic compounds show differences in their dose-response over time.

The Q-KEDRF toolbox

Quantitative methods are often a good way to understand modulating factors. When a sufficient number of experiments determine the procession/cascade of KEs on both dose- and time-scales, quantitative methods are less necessary to obtain an understanding of the MOA. In such cases, the Dose-Time Concordance Table will suffice, and such was the case for DMA.

The relationship of KEs to the critical effect/apical or adverse outcome can be understood by expressing the tumor BMD as a multiple of the BMD values of various KEs (*e.g.*, Simon et al. 2009). BMD_{10} values are shown on the figures in Table 3. Values for the BMD multiple for the three KEs, cytotoxicity, proliferation and hyperplasia, can be determined as:

$$\text{BMD Multiple} = \frac{\text{BMD}_{\text{ApicalEvent}}}{\text{BMD}_{\text{KeyEvent}}} \quad (1)$$

Using Eq. (1), one can determine that the tumor POD is almost 100 fold greater than the BMD_{10} for cytotoxicity at 10 weeks, about 3 fold greater than the BMD_{10} for proliferation at 10 weeks, and about 1.5 fold greater than the BMD_{10} for hyperplasia at 10 weeks. These values provide a means of judging the relative position of the various KEs along the dose continuum.

Quantitative dose-response methods also may prove very useful for understanding and refining proposed MOAs. For example, Simon et al. (2009) used both potency and steepness to determine the dose progression of likely KEs in the MOA for rodent liver tumorigenesis by dioxin. This approach was used again to examine nuclear receptor activation leading to tumor promotion (Budinsky et al. 2014, Corton et al. 2014).

While no single method is appropriate for all situations, the methods described in this section are all part of the Q-KEDRF toolbox. Contrast tests and regression analysis using well-established statistical methods may prove useful for ordering events within a hypothesized MOA (Bretz et al. 2005, Sawilowsky 2002, Tukey et al. 1985). Lutz and Lutz (2009) provide full details of their "hockey stick" model and an R script for ease of use. For developing dose levels corresponding to specified response levels (i.e., benchmark doses), Murrell et al. (1998) suggest the use of the calculated slope of the dose-response and baseline projection. Silkworth et al. (2005) implemented a form of this method but did not describe details of their calculation. The method was fully developed, including

calculation of confidence intervals in Budinsky et al. (2010). Sand et al. (2006) used the second and third derivatives of the dose-response function to obtain a "transition dose range." Further, they identified a response level of 21% as the transition point for the Hill model.

Naïve practitioners may be tempted to use of the numerical results of a single method as a quantitative threshold. In this regard, any quantitative estimate of a threshold needs to be considered in the light of biological significance, and quantitative estimates of thresholds and transitional dose values (TDVs; see Section 4 below) from a variety of methods should be developed (Budinsky et al. 2010). The discussion of thresholds in Slob and Setzer (2014) is particularly enlighting. Notable is their argument that dose is better represented on a logarithmic scale than on a linear one. The use of logarithms with dose is consistent with thermodynamic principles (Waddell 2005, Waddell 2008). This caveat notwithstanding, the ability to obtain quantitative dose values within the low-dose region can greatly help determine the order in dose and time of events within a hypothesized MOA (See Supplementary Content for an example).

Modulating factors—accounting for variation within the human population

The application of the MOA/HRF and the QKEDRF can provide informative and quantitative descriptions of the MOA and dose-response for adverse outcomes (cancer and non-cancer) including those at low, environmentally relevant exposure levels. Such an approach is essentially designed to describe the form of the dose-response curve for a generalized population. What is also needed is an approach that allows for incorporation of the influence of ModFs on the dose-response of KEs that will ultimately enable the quantitative population-level assessment of risk at low exposure levels. ModFs should be understood in terms of their effects on biological processes and KEs within an MOA. The effect of a low protein vegetarian diet on the availability of S-adenosyl methionine as a possible ModF for the toxicity of DMA^V has already been discussed.

One universal ModF is likely to be individual variation in reserve capacities, for example, differing amounts of reduced glutathione that affect the occurrence of particular KEs between individuals and over time within a single individual. Other examples would be the expression of the p53 gene product or the occurrence of oxidative DNA damage.

Variations in the intracellular level of a large number of transcription factors and cofactors can alter both the efficacy and potency for both steroid and glucocorticoid hormones (Blackford et al. 2012, Simons 2010, Sun et al. 2008, Zhang et al. 2012). In fact, limitations in the amount of coregulatory proteins available within the transcription complex may lead to non-monotonic dose response curves such as squelching (Charlier 2009, Kraus et al. 1995, Zhang and Teng 2001). Graphical analysis of these changes yields valuable mechanistic information when the production of the apical response follows a first-order Hill dose-response curve (Dougherty et al. 2012, Ong et al. 2010, Simons and Chow 2012). However, regardless of the order of the dose-response curve of the adverse outcome/apical response, the magnitude and/or position of the dose-response curve will likely be similarly modified by any chemical that binds to nuclear receptors and/or other transcription cofactors.

There may exist many potential ModFs for any particular exposure scenario (*e.g.*, specific chemical, type of exposed individual or group). Therefore, organizing these factors based on common biological mechanisms would be helpful. By doing so, the likelihood of a ModF affecting a particular MOA could be determined. One approach described here is to identify a list of general ModFs that can be broadly separated as Host, Life Style and Environment (Table 1). Other classification schemes for ModFs, perhaps based on MOA, will likely emerge as risk assessment practitioners gain experience with the Q-KEDRF. The OECD is currently developing a program on AOPs, and the International QSAR foundation is developing an "Effectopedia" to provide information about AOPs/MOAs as part of a global scientific collaboration; the Q-KEDRF will likely interface quite well with these efforts (Ankley et al. 2010, Patlewicz et al. 2013). The use of the term "Initial Molecular Event" (IME) to refer to the first step Event, as suggested by Patlewicz et al. (2013), is appropriate and conveys an accurate message—that the initial event may not obligatorily lead to the adverse outcome.

Examples of modulating factors

Two examples are presented below with the goal of improving the understanding of how ModFs can affect KEs and potentially impact the dose-response for the adverse outcome. These examples illustrate different aspects of KEs within biological pathways: xenobiotic processing (metabolism) and endocrine stimulation.

Example 1: Genetic variation in PON1 potentially modulates chlorpyrifos metabolism and toxicity

The MOA for OPs is well known—inhibition of cholinesterases with toxicity manifested as central and peripheral cholinergic effects (Figure 4) (Mileson et al. 1998). Cholinesterase inhibitors include carbamate insecticides, physostigmine used to treat glaucoma, and Δ^9-tetrahydrocannabinol, the active moiety in marijuana. Paraoxonase 1 (PON1) is an arylesterase that metabolizes organophosphate compounds (OPs). Thionophosphorus OPs such as chlorpyrifos (CPF) are metabolized to the oxygen analog or oxon by CYP450 mixed function oxidases. These oxons are potent inhibitors of acetyl cholinesterase (AChE). CPF oxon is inactivated by PON1 in the liver and other tissues (Smith et al. 2011, Timchalk et al. 2002a; 2002b).

Host factors—genetic variability and lifestyle factors. In humans, PON1 activity is age-specific, increasing about 3.5 fold between birth and 7 years of age, remaining constant thereafter (Figure 5) (Smith et al. 2011). Genetic polymorphisms exist in the coding regions of PON1 gene with consequent variation in catalytic activity. For example, PON1 polymorphism at amino acid 192 [glycine (Gln; Q allele) to arginine (Arg; R allele) substitution] changes PON1-mediated esterase activity depending on the substrate present (Adkins et al. 1993). PON1 (R192) hydrolyzes CPF oxon more efficiently than PON1 (Q192) (Richter et al. 2009). Along with the general increase in activity with age, differing phenotypes mature at different rates (Huen et al. 2010). Polymorphisms exist in the

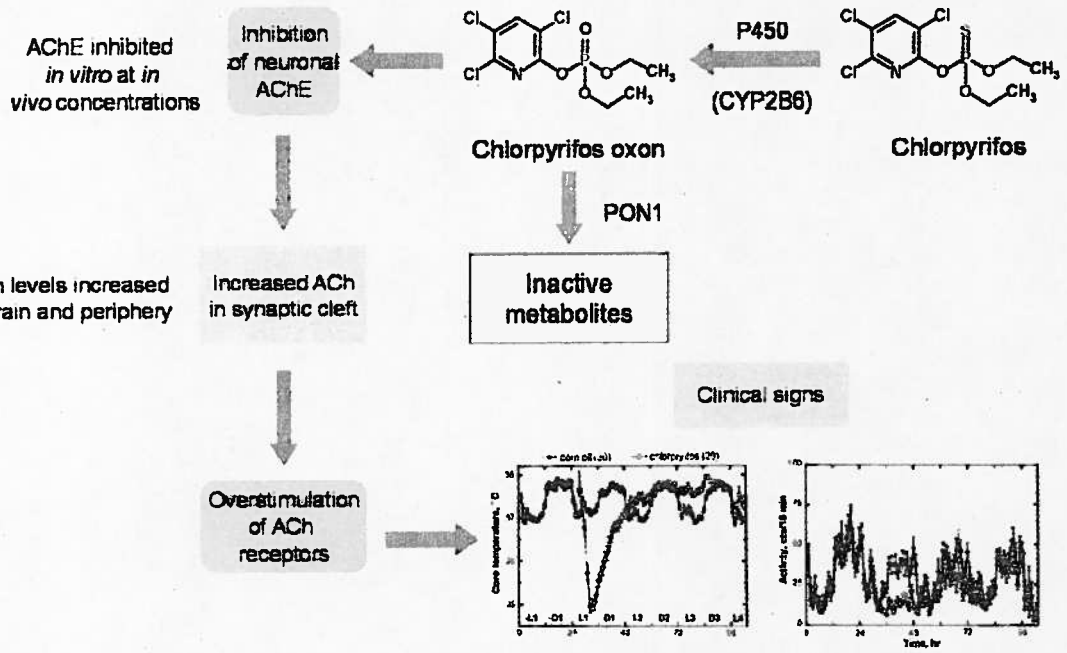


Figure 4. Mode of Action of Chlorpyrifos showing metabolic activation to CPF-oxon and inhibition of acetylcholinesterase as the critical effect. (Figure courtesy of Dr. Alan Boobis).

promoter region of PON1 and may affect expression level and tissue activity. A single nucleotide polymorphism located 108 bases before the transcription start site ($PON1_{-108}$) accounts for 22.4% in the variability in arylesterase activity (Brophy et al. 2001, Deakin and James 2004). Overall, an individual's

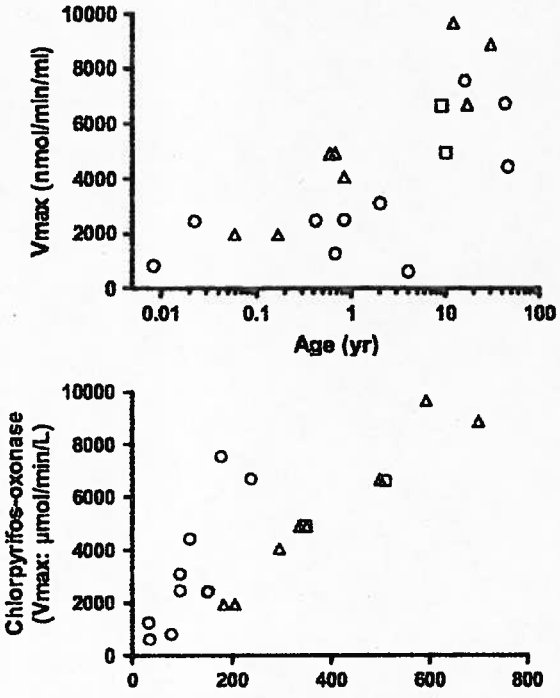


Figure 5. PON1-mediated Vmax values vs. age (upper plot). PON1 functional phenotypes are represented by open circles, open triangles, and open squares for QQ, QR, and RR, respectively (see text for definitions). CPF- oxon hydrolysis V_{max} values in plasma over paraoxon hydrolysis activity (lower plot) resolves QQ and QR, but not QR and RR. (From Smith et al. 2011; permission to reproduce figures granted by Dr. Jordan Smith, 22 March 2013.).

PON1 activity is dependent on variations in the coding region as well as the promoter region. Both the polymorphisms and the age-dependent increase in activity would be categorized as host factors. The age-dependent increase in V_{max} in plasma PON1 activity on a plasma volume basis for individuals of all three genotypes (QQ, QR and RR) is shown in Figure 5.

In addition to these host factors, a number of lifestyle factors affect PON1 activity. Statins are cholesterol-lowering substances that occur naturally in red rice yeast and are also prescribed as drugs. In some human studies, very modest increases in serum PON1 have been observed in those taking statins. However, in other studies, no effect is seen (Costa et al. 2011). Moderate alcohol consumption appears to increase serum PON1 (Sierksma et al. 2002). Pomegranate juice contains several polyphenols and its consumption increases plasma PON1 activity in normal humans and in diabetic patients (Aviram et al. 2000, Rock et al. 2008). The lifestyle factors increase PON1 activity and would tend to desensitize individuals to the effects of thionophosphorus OPs.

Consideration of modulating factors in a chlorpyrifos risk assessment. For risk assessment purposes, the question that must be asked is whether changes in PON1 actually translate into changes in sensitivity, and, if so, whether these host and/or lifestyle factors produce sufficient variation in PON1 activity such that individuals with a sensitive phenotype such as QQ or the very young might constitute an at-risk subpopulation.

When workers exposed to CPF during manufacture were compared to a referent group of chemical workers, no effect of PON1 phenotype was observed (Albers et al. 2010, Garabrant et al. 2009). Urinary 3,5,6-trichloro-2-pyridinol (TCPy) is a metabolite of CPF and a specific biomarker of exposure (Alexander et al. 2006); TCPy levels in all exposed workers were less than those paralleling previously determined no-observed-effect levels for red blood cell (RBC) AChE inhibi-

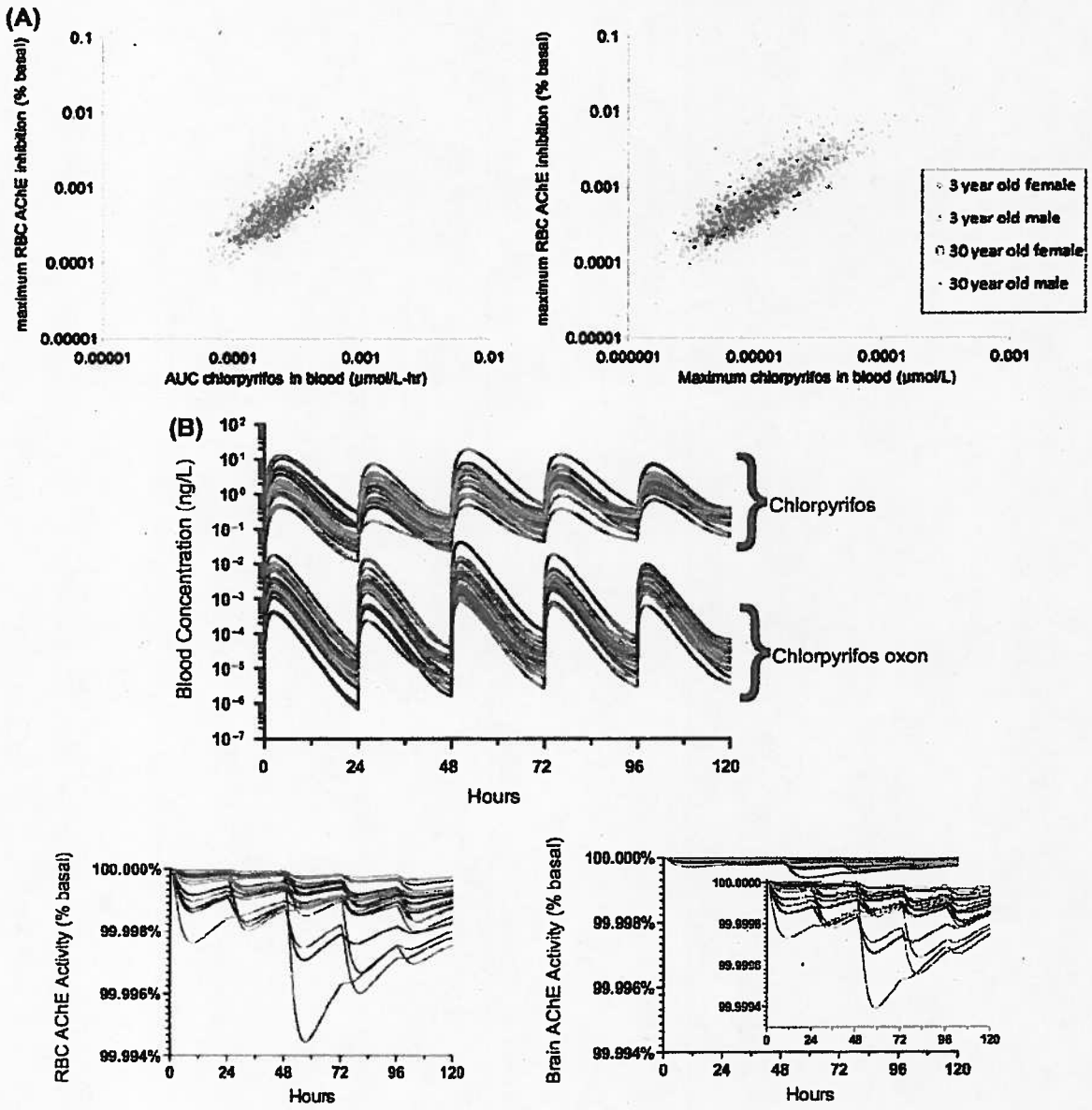


Figure 6. Modeled chlorpyrifos pharmacokinetics in adults and children and resulting AChE inhibition in erythrocytes. A. RBC AChE inhibition from the AUC (left) and maximum CPF concentrations (right) in blood (from Hinderliter et al. 2011). B. Modeled time courses of CPF and CPF oxon in blood from dietary exposures (upper panel) and corresponding RBC AChE inhibition (lower panel). (Reprinted from *Regulatory Toxicology and Pharmacology* (Hinderliter, P.M., Price P.S., Bartels M.J., Timchalk C., Poet T.S. 2011. Development of a source-to-outcome model for dietary exposures to insecticide residues: An example using chlorpyrifos, Regul. Toxicol. Pharmacol. 61, 82–92) with permission from Elsevier.).

tion and changes in neurological function (Albers et al. 2004a; 2004b; 2004c; 2007, 2010, van Gemert et al. 2001).

Enzyme kinetics of PON1 were analyzed in liver microsomes and plasma in both children and adults to measure quantitative age-dependent differences (Smith et al. 2011). These data were incorporated into a probabilistic physiologically based pharmacokinetic and pharmacodynamic (PBPK/PD) model for CPF (Price et al. 2011, Timchalk et al. 2002a, Timchalk et al. 2002b). With this model, the relationship between urinary TCPy and either plasma butyrylcholinesterase (BuChE) or RBC AChE was determined and related to the exposure to CPF. Model results are shown in Figure 6. In three-year-old children, the greatest percent reduction in ChE levels for typical dietary intake was 0.001%. In addition, a sensitivity analysis of the PON1 parameter in blood and liver revealed only a modest influence of this factor. The presence of other detoxification enzymes established a lower limit for the effect of PON1 variation (Hinderliter et al. 2011, Price et al. 2011).

In contrast, at a dose of 300,000 ng/kg/d of CPF, typical of a high-dose animal study, the model indicated that both the age-dependence and the polymorphisms in the activity of hepatic PON1 would be reflected by substantial differences in RBC AChE levels; however, neither these age-dependent differences nor PON1 enzyme polymorphisms are likely to affect RBC AChE levels at real-world human exposure levels (Garabrant et al. 2009, Hinderliter et al. 2011, Smith et al. 2011, Timchalk et al. 2002a; 2002b).

To incorporate ModFs into risk assessment, the effect of these factors needs to be considered at the point of departure or at current exposure levels and not in a purely

DOI 10.3109/10408444.2014.931925

abstract way. An effect of human variation in PON1 on RBC AChE inhibition was observed in the model output at a dose of 300,000 ng/kg/d of CPF but not at current dietary exposures of children and adults for which the respective doses are estimated to be less than 11 ng/kg/d and 3.4 ng/kg/d. Increased sensitivity was not observed at dietary exposures because the exposures were too low to produce a biologically meaningful change in the activity of various cholinesterases, even in sensitive individuals. In addition, individuals of the RR phenotypes appear to have higher activity of PON1 in plasma, thus providing similar capacity for clearance (Figure 5; Smith et al. 2011). Therefore, while the presence of polymorphisms and the age-dependence of PON1 provide illustrations of potential ModFs, the actual effects of these factors must be considered in the context of the entire dose-response curve and relevant exposure levels.

This examination of the MOA for CPF-inhibition of AChE includes tiers 1 through 4 of toxicity resources in the RISK21 roadmap (Figure 1). *In vitro* and *in vivo* data from humans were included; a PBPK/PD model was used for IVIVE and the Q-KEDRF was used to evaluate the ModFs of age and genetic polymorphisms. This probabilistic model is an excellent example of the use of quantitative MOA information in a risk assessment.

Table 4 provides an example of the Species Concordance table for ModFs and presents some of the information discussed above. The table format is sufficiently flexible to accommodate both qualitative and quantitative information. Although the information for CPF was obtained from humans, the columns for animals represent placeholders for those situations in which species extrapolation of the effect of ModFs needs consideration.

Example 2: Factors that can modulate the uterotrophic response

Estrogens induce uterotrophy through activation of the estrogen receptor alpha (ERα), a ligand-activated nuclear receptor and transcription factor. Cellular and physiological factors can modulate the estrogen dose-response for ERα activation, subsequent KEs, and uterine weight gain, the latter considered to be the critical effect in this example. A positive uterotrophic response for a chemical indicates a potential for endocrine disruption (OECD 2003).

Progesterone opposes estrogenic effects and reverses estrogen-induced uterotrophy. Progesterone stops cell growth and prevents the uterine lining from shedding. Like estrogen, progesterone is a ligand that activates a transcription factor. All transcription factors require cofactors for transcription to occur. One function of these cofactors is to increase the activity of RNA polymerase II, sometimes by facilitating chromatin remodeling and RNA polymerase II access to transcriptional start sites. For constitutively expressed genes, chromatin remodeling plays a smaller role than other gene regulatory factors (John et al. 2011). In contrast, RNA polymerase II is already bound at the transcription start site of a large number of other genes and the binding of a transcription factor is the signal for the polymerase to "start" (Levine 2011). Cofactors that interact with both the estrogen and progesterone receptors include steroid receptor coactivator-1 (SRC-1), receptor interacting protein 140 (RIP140), and the histone acetyl transferase chromatin-binding protein/p300 (Kobayashi et al. 2010, Simons 2008, Simons 2010).

Among the mechanisms by which progesterone is proposed to antagonize estrogen actions is by binding to progesterone receptors (PRs) to form complexes that compete with ERαs for cofactors that help mediate and thus increase ERα-mediated gene transactivation (Giangrande et al. 2000, Kraus et al. 1995, Parisi et al. 2009, Wen et al. 1994). In general, the effects of progesterone oppose the effects of estrogen. Thus, the dose-response curve shifts to the right and the system or individual becomes less sensitive to the effects of estrogens. Given that estrogens induce synthesis of PRs, these combined effects may serve as a means of feedback inhibition of estrogen-activated responses.

Uterotrophy as a model system for understanding MOA. Estrogen-induced uterotrophy in rats is an extensively studied response that has been documented to proceed through estrogen binding to the intracellular ERα as the MIE and is a KE in the MOA for the uterotrophic response. The induction of several genes (i.e., ornithine decarboxylase, glucose-6-phosphate dehydrogenase, lactoferrin, c-fos, and uterine peroxidase) occurs in response to estrogen, and these gene expression changes have been proposed as KEs in the MOA of estrogen-induced uterine growth (Figure 7; OECD 2003). Microarray assays have identified various other genes that may also be part of the overall MOA (Henneweer et al. 2007, Naciff et al. 2003).

The effects produced by ModFs shown in Table 5 can modify gene function not only through direct effects on DNA and chromatin but also by altering the strength of the various

Table 4. Dose-response concordance table for Modulating Factors (MFs) in the MOA of chlorpyrifos.

	Qualitative concordance			Quantitative concordance and quantitative Dose-response		
Event or factor	Animals	Humans	Concordance	Strength	Animals	Humans
Modulating Factors MF and affected KE	Animals	Humans	Concordance	Strength	Effects in Animals	Effects in Humans
MF #1 Genetic Polymorphism	NA	R vs. Q allele	NA			QQ genotype more sensitive, but at current exposure levels this difference is not a factor
MF #2 Use of Statin drugs	NA	Statins increase PON1 activity	NA			Statins modestly increase PON1 activity, but the effect is not consistently observed
MF #3 Alcohol Use	NA	Alcohol use increases PON1 activity	NA			This effect is likely not a factor at current exposure levels

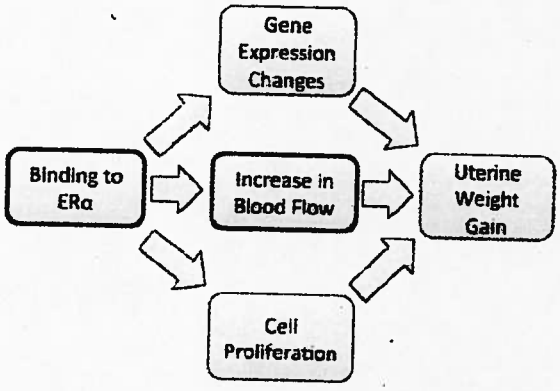


Figure 7. Putative MOA for the uterotrophic response.

binding reactions occurring during gene transcription, including interactions between DNA and protein, between RNA and protein, between DNA and RNA, and between various proteins. The effect of these associations on dose-response is not clear at this time. However, the Q-KEDRF approach allows one to test the prediction that chemicals and factors with similar molecular targets will evoke comparable changes in the adverse outcome/apical event.

The rat uterotrophic response to estrogens was selected for a case study of the utility of using a MOA approach. The first step, of course, was to identify KEs or AEs that could serve as biomarkers for these KEs. Given the abundance of experimental data over the years for rat uterotrophy, this task was expected to be a relatively straightforward application of the new framework (Figure 2). OECD (2003) identifies binding to ERα as the MIE and provides a list of early and late events associated with uterotrophy. Unfortunately, dose-response and timing of these early and late events have not been obtained from the same species or preparation and thus, it is difficult to array these in a meaningful Dose-Time Concordance table. However, guidance from OECD as well as the scientific literature was used as the basis of a putative MOA and a set of proposed KEs for uterotrophy (Figure 7). Given the extent of investment in testing for endocrine effects and the relative maturity of the uterotrophic assay, the lack of information from the same or at least comparable studies seems surprising. This situation emphasizes the need to design studies that address the particular question at hand as it relates to elucidation of the MOA, and illustrates how effective the MOA framework can be in rapidly and effectively identifying critical data gaps. Consideration of MOA as early as possible in the risk assessment process would foster the collection of appropriate data to inform the MOA based on the expected value of the information (Meek et al. 2014a, Meek et al. 2014b). Such an approach would be entirely consistent with the method of problem formulation described in NRC (2009).

Table 5. Cellular effects of modulating factors.

Effect	Sub-effect
Gene Structure	Mutations
	Deletions
	Duplications
Gene Expression	Transcription factors
	Co-activators/accelerators
	Co-repressors/decelerators
	Co-modulators
Post-translational modifications	Acetylation
	Methylation
	Phosphorylation
	Others

Following absorption of estrogen or an estrogenic chemical, binding to ERα would be the MIE. This binding has been measured in a number of species *in vivo* and in cell-free preparations (Levin et al. 1993, Notides et al. 1981). Following receptor binding, early events would include (1) altered expression of estrogen sensitive genes; (2) an increase in uterine blood flow; and (3) an increase in cell proliferation. Respectively, these events can be measured by: (1) microarrays or qRT-PCR; (2) flow transduction or weight gain; and (3) mitotic index or BrdU labeling. Because of the lack of sufficient data from a single high-quality study, as already stated, it is difficult to determine the exact role of these putative KEs in the MOA, but assessing the whole body of evidence using a WoE analysis, KEs can be substantiated. The apical event is, of course, uterine weight gain. At the present time, the order and timing of the changes shown in the third and second columns of Figures 7 and 8, respectively, are not known (Ashby et al. 1999, Gorski et al. 1977, Heneweer et al. 2007, Kaye et al. 1971, Naciff et al. 2003, OECD 2003).

At this point, conclusive identification of putative KEs becomes difficult due to: (1) variations in experimental systems; (2) the absence of data representing multiple KEs from the same study or same laboratory; and (3) and insufficient data points to make quantitative conclusions about dose-response.

Identification of key events for uterotrophy using WoE. Absorption is considered part of absorption, distribution, metabolism, and excretion, and is thus not identified as a KE, although it is the initial event in the process. For some chemicals, metabolic transformation that occurs close in time to absorption may either bioactivate these chemicals to toxic/active metabolites (e.g., polycyclic aromatic hydrocarbons/tamoxifen and cortisone, respectively) or detoxify/inactivate them (e.g., CPF oxon/cortisol) (Chapman et al. 2013, Furr and Jordan 1984). Estrogenic compounds contain one or more phenol groups and, following oral exposure, may be inactivated before reaching the systemic circulation by first-pass phase II metabolism in enterocytes or the liver (e.g., Hengstler et al. 2011). Hence, for estrogenic compounds and uterotrophy, metabolic transformation would not be a KE; however, metabolism may be a KE for other substances that are transformed to toxic metabolites (e.g., dimethylarsinic acid).

For uterotrophy, the MIE of binding to ERα will be a KE if it is empirically observable, and it is very probable that cell proliferation is also a KE. Two KEs can actually be conclusively identified on the basis of counterfactual reasoning and are shown with a thicker outline of the event boxes in Figure 7. The basis for identifying binding to ERα as a KE is the fact that estrogen-receptor knockout mice do not show evidence of cell proliferation, that is, DNA synthesis, in response to estrogen (Curtis et al. 1996, Klotz et al. 2002). However, other responses associated with estrogen-induced uterotrophy such as water imbibition and lactoferrin induction are maintained in the absence of ERα (Das et al. 1997, Winuthayanon et al. 2010). The basis for identifying the increase in blood flow as a KE is the disruption

DOI 10.3109/10408444.2014.931925

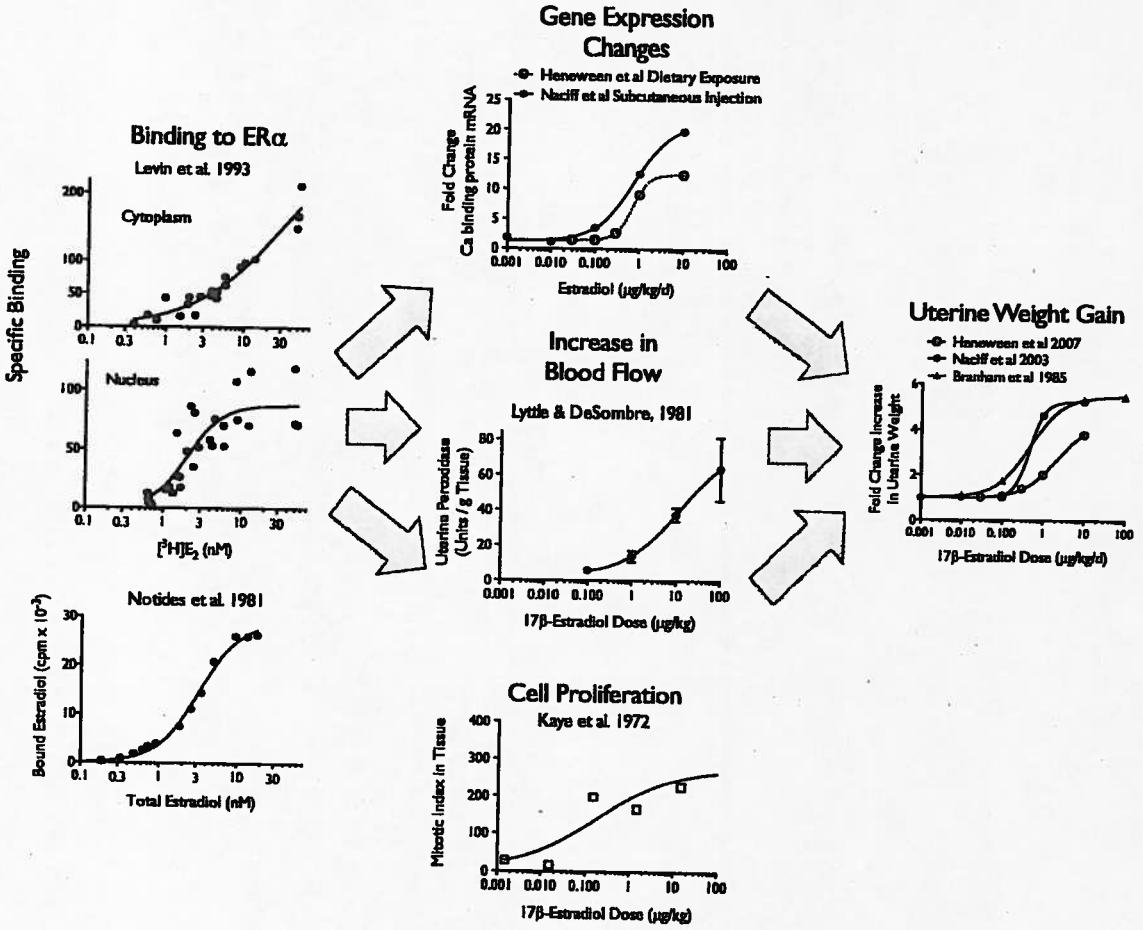


Figure 8. Dose-response plots for putative key events in the MOA for the uterotrophic response.

of the uterotrophic response by L-NG-nitroarginine methyl ester (LNAME) that blocks nitric oxide synthase (Rao et al. 1995, Rosenfeld et al. 1996). Alternatively, the production of catechol estrogens due to an estrogen-mediated increase in peroxidase may also contribute to alpha-adrenergic activation, vasodilation of the uterine arteries, and a consequent increase in blood flow (Lyttle and DeSombre 1977, Farley et al. 1992, Stice et al. 1987a; 1987b). In this example, the increase in uterine peroxidase is being identified as an AE to represent the increase in blood flow (Figure 8).

Dose-response modeling can elucidate the MOA for uterotrophy. Table 6 shows values for Hill model fits for the various responses of KEs and putative KEs. When data are available from a single study, both the EC_{50} and the slope of the dose-response curve are important in understanding the MOA and the relationship to the apical response (e.g., Simon et al. 2009).

EPA's Cancer Guidelines (USEPA 2005a) suggest the possibility of using an earlier KE as a precursor to the apical event and developing a toxicity criterion using the dose-response of this KE. Caution is warranted when using a KE as the basis for development of a toxicity criterion when the dose-response of the KE has a higher value of the Hill coefficient than the apical response; steeper dose-response curves (higher Hill coefficients) will have greater nonlinearity than a first-order Hill response and thus, the rising phase of the dose-response may commence at a higher dose value. Therefore, using the dose-response of the KE as the basis of a toxicity factor may not be a health-protective choice in the case of an apical event or critical effect known to follow a first-order Hill function, as is the case for uterotrophy (OECD 2003). By the same reasoning, the use of an early KE as the basis of a toxicity factor may be inappropriately over-conservative when the KE exhibits a shallower dose-response curve (lower Hill coefficient) than does the critical effect/adverse outcome.

The variation in the Hill coefficients observed in Table 6 is likely a reflection of the fact that these data were obtained from disparate sources. The plots of estrogen binding in the left column of Figure 8 were obtained *in vitro* and thus, IVIVE would be needed to set these on a similar dose scale as whole animal effects.

At this time, most available dose-response curves for estrogen-induced genes and other responses associated with uterotrophy have so few data points that the determination of quantitative aspects of dose-response becomes problematic. Even after all the years of studying uterotrophy, the shape of the curve for the critical effect of uterine weight gain has not been firmly established (Note the variation between the three curves in the rightmost plot of Figure 8).

For these reasons, even the relatively superficial MOA for uterotrophy cannot yet be constructed without new, more detailed data. First, high-quality dose-response curves with more data points for intermediate responses are critical so that an accurate determination of the position (*i.e.*, EC_{50}) and shape

Table 6. Quantitative aspects of the dose-response of key events in the uterotrophic response.

Key event	Study	Hill model parameters	Transition dose values	
			Starting points and slope-based TDVs (Murrell et al., 1998)	BMD_{21} as a transitional dose (Sand et al., 2006)
Binding to ERa	Levin et al. (1993) (cytosol) (fractional binding response)	K_d = 31.2 nM Log K_d = 1.49 n = 0.76	(13.8, 102.8) (5.8, 62.4) **1.26 nM**	**5.41 nM**
	Levin et al. (1993) (nucleus) (fractional binding response	K_d = 2.04 nM Log K_d = 0.310 n = 1.94	(1.49, 27.3) (2.85, 51.8) **1.05 nM**	**1.03 nM**
	Notides et al. (1981)[a] (fractional binding response)	K_d = 3.25 nM Log K_d = 0.512 n = 1.61	(1.853, 7.48) (4.99, 20.7) **1.11 nM**	**1.43 nM**
Gene expression changes in relation to uterine weight gain (Naciff et al., 2003)	Naciff et al. (2003) (fold increase in Ca binding protein)	Bmax = 22.82 fold K_d = 0.807 µg/kg/d Log K_d = -0.0930 n = 0.755	(0.1, 3.5) (1.0, 12.5) **0.082 µg/kg/d**	**0.140 µg/kg/d**
	Naciff et al. (2003) (fold increase in uterine weight)	Bmax = 5.48 fold K_d = 1.26 µg/kg/d Log K_d = 0.010 n = 0.914	(0.1, 1.12) (1.0, 4.63) **0.171 µg/kg/d**	**0.240 µg/kg/d**
Gene expression changes in relation to uterine weight gain (Heneweer et al., 2007)	Heneweer et al. (2007) (fold increase in Ca binding protein)	Bmax = 12.66 fold K_d = 5.35 µg/kg/d Log K_d = 0.728 n = 1.54	(0.3, 2.62) (1.0, 9.1) **0.290 µg/kg/d**	**2.26 µg/kg/d**
	Heneweer et al. (2007) (fold increase)	Bmax = 3.8 fold K_d = 15.65 µg/kg/d Log K_d = 1.195 n = 1.191	(1.0, 2.02) (10.0, 3.79) **0.220 µg/kg/d**	**5.15 µg/kg/d**
Cell proliferation	Kaye et al. (1971) (increase in mitotic index	Bmax = 276 figures K_d = 0.809 µg/kg Log K_d = -0.092 n = 2.21	(1.5, 166) (15, 227) **0.0073 µg/animal/d**	**0.444 µg/animal/d**
Increase in blood flow measure by uterine peroxidase	Lyttle and DeSombre (1977)	Bmax = 69 units/g K_d = 17.3 µg/animal Log K_d = 1.24 N = 0.561	(1.0, 13.8) (10.0, 37.6) **0.053 µg/animal/d**	**1.64 µg/animal/d**
Uterine weight gain	Branham et al. (1985) (mg wet weight)	Bmax = 5.4 fold K_d = 1.85 µg/animal/d Log K_d = 0.268 n = 0.271	(0.1, 1.73) (10.0, 5.31) **0.078 µg/animal/d**	**0.014 µg/animal/d**

Here, the results of two methods for determining transitional dose values (TDVs) are shown. Details of the calculation methods are provided in the text. The form of the Hill model used here is shown in Table 7.

[a]Notides et al. (1981) observe the Hill coefficient for E2 binding to cell-free preparations ERα varies with the concentration of the receptor (from n = 1.1 at 0.3 nM ERα to 1.6 at 3.0 and 4.8 nM ERα, indicating that the Hill coefficient increases with increasing concentrations of ERα.

(*e.g.*, first- or second-order Hill plot) of the curve is possible. This level of information is needed for all events being considered as KEs. These data would be invaluable in eliminating proposed KEs for which the parameters of the dose-response curve are not compatible with those of the apical response. For example, a proximal event that displays a second-order Hill dose-response curve could not be a step in an apical response that exhibits a first-order Hill dose-response curve (Ong et al. 2010, Chow et al. 2011). In this way, quantitative dose-response modeling may provide some mechanistic insights into the role of various events (Simons and Chow 2012). In addition, various analytical tools can be employed to gain mechanistic insight that is available only when the Hill coefficient is equal to one (Dougherty et al. 2012, Ong et al. 2010). A Hill coefficient of two or greater may indicate involvement of transcription factors that act as dimers or higher-order multimers. Alternatively, the observation of Hill coefficients greater than one may also result from ligand-induced conformational changes in binding proteins that function as dimers or multimers (Koshland 1996, Koshland and Hamadani 2002, Levitzki and Koshland 1969). Furthermore, it would be instructive to know the details of ligand binding to ERα in cell-free extracts, in whole cells and in whole animals. One would also want data on the genomic responses *in vitro* and in whole animals. In addition, these data would need to be of sufficient quality to support quantitative dose-response modeling.

Second, additional data are needed to provide dose-response information at different times for those events hypothesized to be KEs. Ideally, these data would be collected under the same experimental conditions as that for the apical event. When performed, interim sacrifices in a cancer bioassay often provide this type of data (*e.g.*, NTP 2006). Such data are necessary for constructing a time line of the KEs and providing data for the Dose-Time Concordance Table (Table 2).

Third, a decision should be made concerning the best experimental system for examining the effects of modulatory factors. For example, if ER-knockout mice are to be used, then high-quality dose-response data, as discussed above, should be collected from

DOI 10.3109/10408444.2014.931925

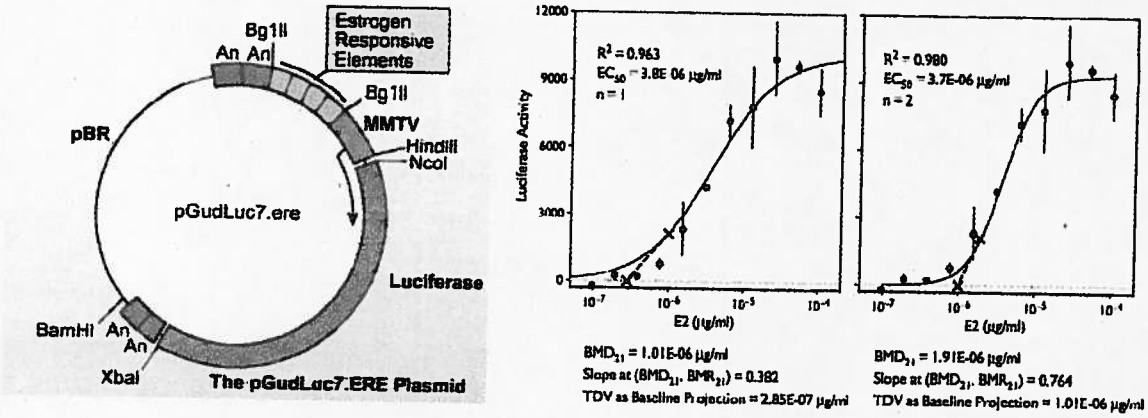


Figure 9. Details of one of the heterologous expression systems that could be used to substitute for the uterotrophic assay. Left: Stably transfected Luc reporter plasmid BG1Luc4E2 cell line from ICCVAM. Right: Concentration-response of the BG1Luc4E2 cells to estradiol showing fits to both first- and second-order Hill functions and the results of the transitional dose value calculation using the baseline projection method (Eq. 3,4 and 5). Please see Supplementary Content for another example.

both normal and knock-out mice. Alternatively, if tissue culture and high throughput studies are selected, then appropriate tissue culture lines could be used and would need to be identified.

Potential utility of understanding the MOA for uterotrophy. One potential result of the greater understanding deriving from more complete experimental data would be the potential for increased usage of in vitro assays measuring KEs and AEs as a screen to identify the chemicals to be assessed further in the uterotrophic assay, a scheme that is consistent with Tox21. The Q-KEDRF seems the best means of demonstrating this consistency. The Interagency Coordinating Committee on the Validation of Alternative Methods has validated a whole cell assay system (Figure 9; BG1Luc ER TA) to assess the activity of different test compounds. Yamasaki et al. (2002, 2003, 2004) measured the response of a reporter gene system as well as the uterotrophic response in whole animals but did not attempt to conduct IVIVE to determine the quantitative relationship between the two—both the reporter gene assay and the in vivo assay were used only for identification of biological effects.

One important aspect of uterotrophy as a model system is that it exemplifies the likely existence of thresholds in MOAs that include receptor binding as a KE. A TDV or range is located at the point where the rising portion of the dose-response begins (Murrell et al. 1998, Sand et al. 2006). Because the binding assays were conducted *in vitro* and the units of dose and routes of exposure were not consistent among the *in vivo* studies, it is difficult to draw conclusions about the numerical values of these either possible threshold values or TDVs, but the ability to estimate these values can, in some cases, provide great insight about the MOA (*e.g.*, Simon et al. 2009).

The value of the Hill coefficient can be important in determining whether linear or nonlinear extrapolation should be used for modeling various KEs or the Adverse Outcome. For the example of uterotrophy here, the ability to obtain insights from quantitative data is mitigated by the relative paucity of the data. Inspection of Figure 9 suggests that for this *in vitro* response in BG1Luc4E2 cells, both first and second order Hill models provide equally good fits to these data. Notides et al. (1981) did observe a shift in the Hill coefficient with increasing concentrations of ERα and attributed this increase to the formation of homodimers with greater availability of ERα. The uterotrophic response itself is generally considered to follow a first-order Hill function but the data from Naciff et al. (2003) seem clearly second order, possibly for this reason. Additional data collection should provide greater certainty regarding the order of the Hill function.

Potential TDVs for the responses in Figures 8 were estimated using the baseline projection method of Murrell et al. (1998) and as the BMD_{21} value as noted by Sand et al. (2006; Table 6). Silkworth et al. (2005) also suggest a method for baseline projection. Details of this method are provided in the next section and in the Supplementary Content.

Alternative Dose Levels from the Hill function for ordering KEs. The Hill model is a three or four parameter equation for a nonlinear relationship between dose and response. The model was first applied by A.V. Hill in 1910 to describe the relationship between oxygen tension and saturation of hemoglobin (Hill 1910). In pharmacology and toxicology, the Hill model has been used extensively to describe the relationship between the dose of a xenobiotic and a biological response (Goutelle et al. 2008, Wagner 1968). In another very recent paper examining the shape and steepness of dose-response relationships for continuous endpoints, the Hill model and the exponential model were both found to provide adequate fits to a large number of data sets covering many continuous endpoints (Slob and Setzer 2014).

For consideration of MOA, location and steepness of the dose-response may help order the events within the dose range. One would wish to know the approximate dose at which the rising portion of the dose-response begins, in other words, the TDV.

A form of the Hill model is shown below and it will be used later to examine responses to estrogenic chemicals. We also provide in Table 7 the inverse equation for calculating the dose at a specified response, for example, the BMD, and the equation for the slope.

$$Response = g + \frac{(V_{max} - g)}{1 + 10^{n(\log_{10}(K_d) - \log_{10}(dose))}} \quad (2)$$

where g = background response;
Vmax = maximal response or efficacy;
n = Hill coefficient (unitless); and

Crit Rev Toxicol, 2014; 44(S3): 17–43

Table 7. Inverse equations and slope equations of dose-response models from EPA's benchmark dose software (USEPA 2012) to enable estimation of baseline projection values.

Model	Equation	Inverse	Derivative
Hill	$\text{Response} = \frac{1}{1+10^{n\left(\log_{10}(K_d)-\log_{10}(dose)\right)}}$	$\log_{10}(dose) = \log_{10}(K_d) - \frac{\log_{10}\left(\frac{1}{\text{Response}}-1\right)}{n}$	$\text{Slope} = \frac{n10^{n\left(\log_{10}(K_d)-\log_{10}(dose)\right)}}{dose\left(10^{n\left(\log_{10}(K_d)-\log_{10}(dose)\right)}+1\right)^2}$
Logistic	$\text{Response} = \frac{1}{1+e^{-(\alpha+\beta dose)}}$	$dose = \frac{-\alpha-\log_e\left(\frac{1}{\text{Response}}-1\right)}{\beta}$	$\text{Slope} = \frac{\beta e^{-(\alpha+\beta dose)}}{\left(1+e^{-(\alpha+\beta dose)}\right)^2}$
Log-Logistic or Dichotomous Hill	$\text{Response} = \frac{1}{1+e^{-(\alpha+\beta\log_e(dose))}}$	$\log_e(dose) = \frac{-\alpha-\log_e\left(\frac{1}{\text{Response}}-1\right)}{\beta}$	$\text{Slope} = \frac{\beta e^{-(\alpha+\beta\log_e(dose))}}{dose\left(1+e^{-(\alpha+\beta\log_e(dose))}\right)}$
Multistage (2nd order)	$\text{Response} = 1-e^{-\beta_1 dose-\beta_2 dose^2}$	$dose = \frac{-\beta_1+\sqrt{\beta_1^2-4\beta_2\log_e(1-\text{Response})}}{2\beta_2}$	$\text{Slope} = e^{-\beta_1 dose-\beta_2 dose^2}(\beta_1+2\beta_2 dose)$
Weibull	$\text{Response} = 1-e^{-\beta dose^{\alpha}}$	$\log_e(dose) = \frac{\log_e\left(\frac{-\log_e(1-\text{Response})}{\beta}\right)}{\alpha}$	$\text{Slope} = \alpha\beta\, dose^{\alpha-1}e^{-\beta dose^{\alpha}}$
Exponential Model 2	$\text{Response} = \alpha e^{\beta dose}$	$dose = \frac{\log_e\left(\frac{\text{Response}}{\alpha}\right)}{\beta}$	$\text{Slope} = \alpha\beta e^{\beta dose}$
Exponential Model 3	$\text{Response} = \alpha e^{\beta dose^{\delta}}$	$\log_e(dose) = \log_e\left(\frac{\log_e\left(\frac{\text{Response}}{\alpha}\right)}{\beta}\right)\left(\frac{1}{\delta}\right)$	$\text{Slope} = \alpha\beta e^{\beta dose^{\delta-1}}e^{\beta dose^{\delta}}$

These values may be useful for ordering events within a hypothesized MOA. These equations are written to be easy to implement in spreadsheet software such as MS-Excel. Their use is not for development of regulatory criteria but rather exploration of hypothesized MOAs.

K_d = affinity or dose at the half-maximal response, a measure of potency (For concentrations, this parameter is often shown as EC_{50}, indicating a dose or concentration with a 50% of maximal efficacy.

In Eq. (2) and all equations following, common or base 10 logarithms are denoted by "$\log_{10}$" and natural logarithms are denoted by "$\log_e$" All the responses shown in Figure 8 were fit to Eq. (2). The third column in Table 6 shows the fitted values for K_d and n, the Hill coefficient.

Another method to obtain the TDV is that of Murrell et al. (1998). The baseline projection of the rising part of the curve is obtained by choosing two points by inspection, one above and one below the half-maximal response. The slope of the rising portion is calculated as the ratio of the differences of the dose and response values of these two points.

$$\text{Slope} = \frac{R_1 - R_2}{\log_{10}(dose_1) - \log_{10}(dose_2)} \quad (3)$$

where R_i = fractional response levels above and below 0.5.

This slope will likely be very close to that at the half-maximal response. Hence, using 0.5 as the measure of the response at the K_d value on a zero-to-one scale, the dose at the onset of the rising portion of the dose-response is calculated as:

$$\text{TDV} = \log_{10}(K_d) - \frac{0.5}{Slope} \quad (4)$$

The results are shown in column 4 of Table 6.

For the form of the Hill model shown in Eq. (2), the dose at any fractional response level, for example, 0–1, can be obtained as follows:

$$\log_{10}(dose) = \log_{10}(K_d) - \frac{\log_{10}\left(\frac{1}{\text{Response}}-1\right)}{n} \quad (5)$$

Equation (5) was used to calculate the BMD_{21}, identified as a TDV by Sand et al. (2006; Table 6).

Once the Hill model parameters for the dose-response (Eq. 2) have been obtained from fitting software, the results of Eqs. (3–5) can be easily obtained with spreadsheet software or even a hand calculator. Only the Hill coefficient, n, and the common logarithm of the half-maximal concentration, $\log_{10}(K_d)$, are needed.

These doses are referred to as transitional because their location marks the approximate transition to the rising portion of the dose-response (Sand et al. 2006). The method of Murrell et al. (1998) explicitly considers steepness with a calculation of the slope. The BMD_{21} is the point at which the generalized Hill model transitions to the rising phase, as indicated by higher derivatives of the model (Sand et al. 2006).

DOI 10.3109/10408444.2014.931925

Measurements of binding to the estrogen receptor show very similar slope-based TDVs. One might expect gene expression changes to occur at a lower dose than uterine weight gain. The slope-based TDV for the increase in expression of vitamin D-dependent intestinal calcium-binding protein (Calb3) from Naciff et al. (2003) is about half than that for uterine weight gain in the same study; however, the BMD_{21} values for these two effects are much more similar (Table 6).

In contrast, the data from Heneweer et al. (2007) show about a two-fold increase in the BMD_{21} but similar slope-based TDVs. Both studies used immature female Sprague-Dawley rats so the difference in the relationship of TDVs between the two studies is likely due to the small number of data points and uncertainty in the fit. The fact that these two methods of calculating a transitional dose range/value give different results for two similar studies would be a reason to obtain further details of the biological role of Calb3 in the uterotrophic response. Highlighting the need for additional qualitative information about the biology underlying the MOA is a great benefit of the use of the Q-KEDRF.

Confidence limits could be likely determined for these TDVs, but the point of their use is to obtain evidence regarding the timing and role of events in a hypothesized MOA. The relationship between Calb3 and uterine weight is not yet known (Naciff et al. 2003, Heneweer et al. 2007). Hence, a review of the literature and possibly some laboratory studies would go further in addressing this particular data gap.

Last in the table are three measurements for increases in uterine cell proliferation, blood flow, and weight gain reported in OECD (2003). All three studies were conducted in rats and the TDVs may suggest that the order of events along the dose continuum is:

1) cell proliferation;
2) increased blood flow measured by uterine peroxidase; and,
3) uterine weight gain.

Both types of TDV for all three studies were expressed in units of μg/animal/d. Here, the slope-based TDV suggests that cell proliferation may be a low dose–response, whereas the slope-based TDVs for increases in blood flow and uterine weight gain occur fairly close to each other along the dose continuum. The TDVs as the BMD_{21} for these three responses are more challenging to interpret. The reason is likely that the slope-based TDVs used the actual data to develop a slope value and the BMD_{21} TDV uses the fitted Hill coefficient. In all three cases, the fitted Hill coefficients had low values and the fits were performed on data with six or fewer dose values.

Another example of this type of quantitative MOA analysis can be found in recent work on the MOA of dioxin liver carcinogenesis in rats (Budinsky et al. 2014, Simon et al. 2009). Both papers present figures showing dose-response plots of different events in the MOA ordered by increasing K_d values and increasing Hill coefficients.

In all likelihood, statisticians can think of much more sophisticated analyses using the slope of the dose-response. Such approaches could use expressions for the slope of the dose-response and attempt to discover in what dose ranges the most rapid change occurs. However, for the purposes of working out events within a hypothesized MOA, easily calculated values such as K_d or the TDV can be very useful.

There may be additional insight gained from using a baseline projection method similar to that obtained at the half-maximal response level using the procedure of Murrell et al. (1998). Table 7 provides equations for commonly used empirical dose-response models, the corresponding inverse equations that solve for dose as the independent variable based on a chosen response, and equations for the dose-response slope at any point. In some instances, these equations can be used to project to the baseline or zero response using the slope at the chosen level of response (Figure 9; Supplementary Content). The inverse equations in Table 7 simply express the dose corresponding to a chosen fractional response (assuming "1" is the maximal response). Using these equations should prove simpler than obtaining an implicit solution. The slope equations in Table 7 provide a means of calculating the slope at the benchmark point (BMD, BMR).

Baseline projection from the 21% response level is shown graphically in Figure 9. Although the values for the EC_{50} are very close, the BMD_{21} values differ by a factor of 2 and the baseline projections from the 21% response level differ by over three-fold. An examination of these differences may help discover the sequence of KEs in a proposed MOA.

As noted, the Supplementary Content provides another example calculation of this baseline projection method that incorporates both the location and steepness of the dose-response at a chosen point and how to use such information in thinking about a hypothesized MOA.

Comparing the values of the Hill coefficients of various events in a hypothesized MOA may provide additional insight and contribute to the decision of whether to assume the adverse outcome follows a linear or nonlinear MOA. Ligand binding and the constellation of early steps in gene transcription may have Hill coefficients close to unity and thus their dose-response might be considered linear (Murrell et al. 1998, Budinsky et al. 2014). KEs that have Hill coefficients with values of 2 or greater invariably indicate the MOA for the adverse outcome will be nonlinear (Chow et al. 2011).

Log-steepness, measured by the ratio of the BMD_{10} to the BMD_{05}, was considered for use in ordering events with a hypothesized MOA (Slob and Setzer 2014). The dose-response data provided in EPA (2005c) was used to obtain values of log-steepness for KEs in the MOA of cacodylic acid (Tables 2 and 3; Figure 3). The three KEs are cytotoxicity, proliferation, and hyperplasia occurring at 10 weeks (Table 3). Appendix D of this EPA publication contains the BMDS output for these three KEs. The values for log-steepness calculated as the BMD ratio for these three KEs (cytotoxicity, proliferation, and hyperplasia) were 2.1, 1.1, and 1.4, respectively. Slob and Setzer (2014) note that log-steepness estimated as the BMD ratio is imprecise, and, while this is only a single example, this easily calculated value did not prove helpful in ordering KEs within a hypothesized MOA. Further work is needed to determine whether this measure of log-steepness can indeed help inform details of MOA.

Constructing a Dose-Time Concordance Table may also help to identify late occurring KEs. These late KEs in the modes of action of complex adverse outcomes such as cancer or developmental effects, may be highly nonlinear and will likely have high-valued Hill coefficients (Brown et al. 2012, Hanahan and Weinberg 2000, 2011, Simon et al.

2009). In some cases, sufficient information about the MOA will be available to select some KEs to use as appropriate precursors to the adverse outcome such as was done by EPA for dimethylarsinic acid. The ability to select appropriate precursor KEs will require quantitative knowledge of the relationship between that KE and the adverse outcome. When the knowledge is available, such precursor events can be used as the basis for risk assessment (Simon et al. 2009, USEPA 2005a, Thompson et al. 2014).

Application of knowledge of the MOA for uterotrophy in risk assessment. A number of host, life stage, and environmental factors likely will modulate human responses to chemicals shown to be estrogenic in the uterotrophic assay and in surrogate in vitro assays. Because many potentially estrogenic chemicals contain one or more hydroxyl groups that interact with specific ligand-binding pockets in ERα, the metabolism of these chemicals in the enterocytes lining the gastrointestinal tract and the liver may result in their inactivation. Hence, for some chemicals, first pass serves as a detoxification process.

For example, bisphenol A (BPA) is almost completely inactivated by phase II metabolism in enterocytes and liver by both glucuronidation and sulfation. These processes occur in both humans and rats (Hengstler et al. 2011). Differences in glucuronidation and sulfation of BPA in rats and humans exist and may provide the basis for interspecies extrapolation of metabolism and consequent bioavailability of BPA (Mazur et al. 2010). Alternatively, these data may be used to improve PBPK models of BPA (Fisher et al. 2011, Teeguarden et al. 2005).

Modulating factors for estrogenic responses in humans. After oral ingestion, it is not possible to detect free BPA in plasma in adult humans (Willhite et al. 2008). PBPK modeling suggests that levels of free BPA in very young children may be higher than in adults due to lower glucuronidation capacity during the first 2 months of life (Edginton and Ritter 2009, Mielke and Gundert-Remy 2009). Free BPA has been detected in the urine of premature infants in neonatal intensive care and its source may be medical devices and the need to deliver medical interventions directly via the blood (Calafat et al. 2009). In contrast, free BPA has not been detected in the urine of full-term healthy infants up to 44 days in age (Nachman et al. 2013). This fact suggests that the glucuronidation capacity in healthy infants is sufficient to metabolize BPA from environmental exposures.

Polymorphisms in uridine 5'-diphospho-glucuronosyltransferase enzymes that conjugate glucuronide may be a potential ModF (Allegaert et al. 2009, Court 2010, Girard et al. 2007, Guillemette et al. 2010, Krekels et al. 2012, Mercke Odeberg et al. 2006, Miyagi and Collier 2011, Strassburg et al. 1997, de Wildt et al. 1999). As noted, differences in glucuronidation occur with gender and age. Diet may also be a factor in the ability to inactivate estrogenic chemicals (Navarro et al. 2009, 2011, Saracino et al. 2009). In all cases of oral exposure, the actual exposure needs to be considered in a quantitative fashion—the inability to detect free BPA in the urine of normal infants suggests that exposures may be sufficiently low that glucuronidation is essentially complete (*e.g.*, Ye et al. 2012). There may be exposures to estrogenic chemicals by routes other than oral, for example, dermal or inhalation, for which glucuronidation does not occur. However, these exposures appear to be miniscule (Geens et al. 2012).

The occurrence of male reproductive tract pathologies in offspring of women administered diethylstilbestrol (DES) during pregnancy suggests that both a lowest-observed-adverse-effect level and a NOAEL exist for these developmental effects. Because no formal clinical trials had been conducted with DES, the total dose varied among clinics by an order of magnitude or more. Male reproductive tract abnormalities were observed in offspring of mothers receiving higher total doses of DES, that is, 12–18 g during pregnancy (Dietrich 2010, Golden et al. 1998), whereas no clear increase was observed in reproductive tract effects in offspring of mothers administered 1.4 g of DES during pregnancy (Leary et al. 1984).

Exposure to more than one estrogenic chemical, such as dietary phytoestrogens, may interact with, or complement, endogenous or other exogenous chemicals. As noted, at sufficient doses, estrogenic chemicals act as anti-androgens in males. However, dose addition of these chemicals is unlikely unless at least two of the doses occur in the rising portion of the dose–response curve (Borgert et al. 2012). Quantitative aspects of dose-response such as affinity, efficacy, and potency need to be considered for chemicals that act via receptor binding—simply using dose addition and some measure of relative potency will be inadequate for risk assessment (Borgert et al. 2012).

The examination of the MOA for uterotrophy requires *in vivo* measurement of the adverse outcome/apical endpoint and includes *in vitro* measurements of the MIE, genomic data, and physiological measures of KEs. Hence, this example demonstrates the use of data from tiers 1–4 of the toxicity resource pyramid of the RISK21 Roadmap (Figure 1), and illustrates the strength of MOA analyses in terms of generating data useful for risk assessment purposes.

Discussion

The MOA/HRF along with the Q-KEDRF described here provides a strong foundation for using the information gathered as a means of reducing uncertainty in risk assessments. The KEDRF laid out the approach for harnessing the extensive available data for the KEs within a putative MOA. The Q-KEDRF provides additional tools with which to gain further insights about how the KEs relate to each other and to the adverse outcome/apical event in a quantitative way in both the dose- and time-dimensions.

In risk assessment, the greatest quantitative impact comes from the choice of a linear approach versus a nonlinear approach for modeling the dose-response for the critical effect or apical effect of concern. The dose-responses for the KEs can be used to inform the shape of the dose-response for the apical effect of concern. For receptor-mediated effects, as noted, quantitative dose-response modeling can provide much greater understanding. For example, if the dose-responses of some or all KEs exhibit biological thresholds, for example, cytotoxicity of the liver and kidney induced by chloroform (Andersen et al. 2000), then the combination of events will also display a dose threshold. Alternatively, if the dose-responses for KEs do not exhibit dose thresholds, then the combination of events may result in a linear dose-response for the apical event. The ability to calculate possible threshold or transition dose values from quantitative dose-response modeling provides a means to determine whether linear or nonlinear extrapolation is appropriate (Table 7).

DOI 10.3109/10408444.2014.931925

It is increasingly clear that account has to be taken of those ModFs that could influence the shape of the dose–response curve, the efficacy or magnitude of the apical response or selected critical effect, and the potency or location along the dose continuum. For example, how much variation can be expected for a particular ModF? Again, this depends on the underlying biology. Sufficient variation may "linearize" the dose-response of the apical event (Conolly et al. 2005, Lutz 2001). The question then is: will this amount of variation "linearize" the population dose-response to a sufficient extent to support the choice of linear low-dose extrapolation? As a generalization, ModFs that are likely to modify the dose-response characterization as part of the risk assessment process will be relatively frequent in the population (given that dose-response is a population feature). Some of these ModFs are "inevitable" and are characteristics of the general population (sex, age, and genotype); others are "manageable" and are characteristic of specific subpopulations (smoking, diet, and weight). Additional research on this topic and the overall role of ModFs is essential to inform the consideration of ModFs and their effect on MOA as part of problem formulation.

At this point in the history of risk assessment, the utility of the Q-KEDRF remains to be determined: experience in conducting real-world risk assessments will demonstrate any value added. Certainly, the Dose-Time Concordance table and Dose-Response Species Concordance table for KEs and ModFs (Tables 2–4) should provide a significant amount of help. The National Research Council recently reviewed EPA's Formaldehyde risk assessment and as part of that review, suggested that the documentation for chemical-specific risk assessments in the IRIS program be organized around informative tables (NRC 2011). The Dose-Time and Dose-Response Species Concordance tables could be very useful in that effort.

At present, the full utility of the Q-KEDRF has barely begun to be realized. The example of rat uterotrophy, while being arguably the best documented physiological response to the extensively studied steroid hormones, clearly demonstrates not only the shortcomings in the available data but also how much actual insight can be acquired through the development of a Q-KEDRF for a specific response. The Q-KEDRF will likely change as experience in using it is gained. Nonetheless, some of the basic issues discussed here will likely become hallmarks of any framework implemented to understand the MOA of a particular adverse outcome. These issues include: (1) separating KEs from putative KEs and (2) understanding the relationship between KEs based upon their dose-response and the timing of their occurrence. Use of this information can significantly improve risk assessments by reducing uncertainty and fostering the incorporation of this information into easy-to-use tables.

Acknowledgments

The authors gratefully acknowledge the government, academic, and industry scientists of the HESI RISK21 Dose-Response Subteam and the HESI RISK21 Steering Team for their contributions to this work. (For a full list of RISK21 participants, see http://www.risk21.org). In particular, the authors thank Dr. Bette Meek (University of Ottawa, Canada) for her critical review of the manuscript. We also thank Dr. Carson Chow (National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health) for his helpful and insightful comments.

Declaration of interest

This manuscript was prepared under the auspices of the International Life Sciences Institute Health and Environmental Sciences Institute (ILSI HESI), a non-profit organization aimed at engaging scientists from academia, government, industry, research institutes, and NGOs to identify and resolve global health and environmental issues. ILSI HESI collects funding from member companies to support projects. The authors' affiliations are as shown on the cover page. The authors had sole responsibility for the writing and content of the paper. The views and opinions expressed in the paper are those of the authors, and do not necessarily reflect the views of the authors' employers or the opinions or policies of the US EPA or NIH. Mention of trade names does not constitute endorsement. None of the authors has recently or is currently involved as an expert witness in litigation or formal government rule-making on the subject of this paper. The authors employed by ILSI HESI participated as part of their normal employment. No other authors received financial support or an honorarium in the preparation of this paper. The authors declare that there are no conflicts of interest.

References

Adami HO, Berry SC, Breckenridge CB, Smith LL, Swenberg JA, Trichopoulos D, et al. (2011). Toxicology and epidemiology: improving the science with a framework for combining toxicological and epidemiological evidence to establish causal inference. Toxicol Sci, 122, 223–34.

Adeleye Y, Andersen M, Clewell R, Davies M, Dent M, Edwards S, et al. (2014). Implementing toxicity testing in the 21st century (TT21C): making safety decisions using toxicity pathways, and progress in a prototype risk assessment. Toxicology, pii: S0300-483X(14)00032-8. doi: 10.1016/j.tox.2014.02.007. [Epub ahead of print].

Adkins S, Gan KN, Mody M, La Du BN. (1993). Molecular basis for the polymorphic forms of human serum paraoxonase/arylesterase: glutamine or arginine at position 191, for the respective A or B allozymes. Am J Hum Genet, 52, 598–608.

Albers JW, Berent S, Garabrant DH, Giordani B, Schweitzer SJ, Garrison RP, Richardson RJ. (2004a). The effects of occupational exposure to chlorpyrifos on the neurologic examination of central nervous system function: a prospective cohort study. J Occup Environ Med, 46, 367–78.

Albers JW, Garabrant DH, Berent S, Richardson RJ. (2010). Paraoxonase status and plasma butyrylcholinesterase activity in chlorpyrifos manufacturing workers. J Expo Sci Environ Epidemiol, 20, 79–89.

Albers JW, Garabrant DH, Mattsson JL, Burns CJ, Cohen SS, Sima C, et al. (2007). Dose-effect analyses of occupational chlorpyrifos exposure and peripheral nerve electrophysiology. Toxicol Sci, 97, 196–204.

Albers JW, Garabrant DH, Schweitzer S, Garrison RP, Richardson RJ, Berent S. (2004b). Absence of sensory neuropathy among workers with occupational exposure to chlorpyrifos. Muscle Nerve, 29, 677–86.

Albers JW, Garabrant DH, Schweitzer SJ, Garrison RP, Richardson RJ, Berent S. (2004c). The effects of occupational exposure to chlorpyrifos on the peripheral nervous system: a prospective cohort study. Occup Environ Med, 61, 201–11.

Alexander BH, Burns CJ, Bartels MJ, Acquavella JF, Mandel JS, Gustin C, Baker BA. (2006). Chlorpyrifos exposure in farm families: results from the farm family exposure study. J Expo Sci Environ Epidemiol, 16, 447–56.

Allegaert K, Vanhaesebrouck S, Verbesselt R, van den Anker JN. (2009). In vivo glucuronidation activity of drugs in neonates: extensive interindividual variability despite their young age. Ther Drug Monit, 31, 411–15.

Andersen ME, Meek E, Boorman GA, Brusick DJ, Cohen SM, Dragan YP, et al. (2000). Lessons learned in applying the U.S. EPA

proposed cancer guidelines to specific compounds. Toxicol Sci, 53, 159–72.

Andersen ME, Preston RJ, Maier A, Willis AM, Patterson J. (2014). Dose-response approaches for nuclear receptor-mediated modes of action for liver carcinogenicity: results of a workshop. Crit Rev Toxicol, 44, 50–63.

Ankley GT, Bennett RS, Erickson RJ, Hoff DJ, Hornung MW, Johnson RD, et al. (2010). Adverse outcome pathways: a conceptual framework to support ecotoxicology research and risk assessment. Environ Toxicol Chem, 29, 730–41.

Ashby J, Tinwell H, Soames A, Foster J. (1999). Induction of hyperplasia and increased DNA content in the uterus of immature rats exposed to coumestrol. Environ Health Perspect, 107, 819–22.

Aviram M, Dornfeld L, Rosenblat M, Volkova N, Kaplan M, Coleman R, et al. (2000). Pomegranate juice consumption reduces oxidative stress, atherogenic modifications to LDL, and platelet aggregation: studies in humans and in atherosclerotic apolipoprotein E-deficient mice. Am J Clin Nutr, 71, 1062–76.

Beyer LA, Beck BD, Lewandowski TA. (2011). Historical perspective on the use of animal bioassays to predict carcinogenicity: evolution in design and recognition of utility. Crit Rev Toxicol, 41, 321–38.

Blackford JA Jr, Guo C, Zhu R, Dougherty EJ, Chow CC, Simons SS Jr. (2012). Identification of location and kinetically defined mechanism of cofactors and reporter genes in the cascade of steroid-regulated transactivation. J Biol Chem, 287, 40982–95.

Boekelheide K, Andersen ME. (2010). A mechanistic redefinition of adverse effects – a key step in the toxicity testing paradigm shift. ALTEX, 27, 243–52.

Boekelheide K, Campion SN. (2010). Toxicity testing in the 21st century: using the new toxicity testing paradigm to create a taxonomy of adverse effects. Toxicol Sci, 114, 20–4.

Boobis AR, Cohen SM, Dellarco V, McGregor D, Meek ME, Vickers C, et al. (2006). IPCS framework for analyzing the relevance of a cancer mode of action for humans. Crit Rev Toxicol, 36, 781–92.

Boobis AR, Daston GP, Preston RJ, Olin SS. (2009). Application of key events analysis to chemical carcinogens and noncarcinogens. Crit Rev Food Sci Nutr, 49, 690–707.

Boobis AR, Doe JE, Heinrich-Hirsch B, Meek ME, Munn S, Ruchirawat M, et al. (2008). IPCS framework for analyzing the relevance of a noncancer mode of action for humans. Crit Rev Toxicol, 38, 87–96.

Borgert CJ, Sargent EV, Casella G, Dietrich DR, McCarty LS, Golden RJ. (2012). The human relevant potency threshold: reducing uncertainty by human calibration of cumulative risk assessments. Regul Toxicol Pharmacol, 62, 313–28.

Bretz F, Pinheiro JC, Branson M. (2005). Combining multiple comparisons and modeling techniques in dose-response studies. Biometrics, 61, 738–48.

Briggs K, Cases M, Heard DJ, Pastor M, Pognan F, Sanz F, et al. (2012). Inroads to predict in vivo toxicology-an introduction to the eTOX project. Int J Mol Sci, 13, 3820–46.

Brophy VH, Jampsa RJ, Clendenning JB, McKinstry LA, Jarvik GP, Furlong CE. (2001). Effects of 5' regulatory-region polymorphisms on paraoxonase-gene (PON1) expression. Am J Hum Genet, 68, 1428–36.

Brown FS, Jacobs A, Fitzpatrick S. (2012). Reproductive and developmental toxicity testing: from in vivo to in vitro. ALTEX, 29, 333–9.

Buchanan RL, Havelaar AH, Smith MA, Whiting RC, Julien E. (2009). The key events dose-response framework: its potential for application to foodborne pathogenic microorganisms. Crit Rev Food Sci Nutr, 49, 718–28.

Budinsky RA, LeCluyse EL, Ferguson SS, Rowlands JC, Simon T. (2010). Human and rat primary hepatocyte CYP1A1 and 1A2 induction with 2,3,7,8-tetrachlorodibenzo-p-dioxin, 2,3,7,8-tetrachlorodibenzofuran, and 2,3,4,7,8-pentachlorodibenzofuran. Toxicol Sci, 118, 224–35.

Budinsky RA, Schrenk D, Simon T, van den Berg M, Reichard JF, Silkworth JB, et al. (2014). Mode of action and dose-response framework analysis for receptor-mediated toxicity: the aryl hydrocarbon receptor as a case study. Crit Rev Toxicol, 44, 83–119.

Calafat AM, Weuve J, Ye X, Hu H, Ringer S, Huttner K, Hauser R. (2009). Exposure to bisphenol A and other phenols in neonatal intensive care unit premature infants. Environ Health Perspect, 117, 639–44.

Chapman K, Holmes M, Seckl J. (2013). 11β-hydroxysteroid dehydrogenases: intracellular gate-keepers of tissue glucocorticoid action. Physiol Rev, 93, 1139–206.

Charlier TD. (2009). Importance of steroid receptor coactivators in the modulation of steroid action on brain and behavior. Psychoneuroendocrino, 34, S20–9.

Chow CC, Ong KM, Dougherty EJ, Simons SS Jr. (2011). Inferring mechanisms from dose-response curves. Methods Enzymol, 487, 465–83.

Clark JH, Markaverich BM. (1984). The agonistic and antagonistic actions of estriol. J Steroid Biochem, 20, 1005–13.

Cohen SM. (1995). Role of urinary physiology and chemistry in bladder carcinogenesis. Food Chem Toxicol, 33, 715–30.

Cohen SM, Arnold LL. (2011). Chemical carcinogenesis. Toxicol Sci, 120, S76–92.

Cohen SM, Arnold LL, Lewis AS, Beck BD. (2006). Methylated arsenicals: the implications of metabolism and carcinogenicity studies in rodents to human risk assessment. Crit Rev Toxicol, 36, 99–133.

Cohen SM, Ellwein LB. (1990). Cell proliferation in carcinogenesis. Science, 249, 1007–11.

Cohen SM, Klaunig J, Meek ME, Hill RN, Pastoor T, Lehman-Mckeeman L, et al. (2004). Evaluating the human relevance of chemically induced animal tumors. Toxicol Sci, 78, 181–6.

Cohen SM, Meek ME, Klaunig JE, Patton DE, Fenner-Crisp PA. (2003). The human relevance of information on carcinogenic modes of action: overview. Crit Rev Toxicol, 33, 581–9.

Conolly RB, Gaylor DW, Lutz WK. (2005). Population variability in biological adaptive responses to DNA damage and the shapes of carcinogen dose-response curves. Toxicol Appl Pharmacol, 207, 570–5.

Corton JC, Cunningham ML, Hummer BT, Lau C, Meek B, Peters JM, et al. (2014). Mode of action framework analysis for receptor-mediated toxicity: the peroxisome proliferator-activated receptor alpha (PPARα) as a case study. Crit Rev Toxicol, 44, 1–49.

Costa LG, Giordano G, Furlong CE. (2011). Pharmacological and dietary modulators of paraoxonase 1 (PON1) activity and expression: the hunt goes on. Biochem Pharmacol, 81, 337–44.

Court MH. (2010). Interindividual variability in hepatic drug glucuronidation: studies into the role of age, sex, enzyme inducers, and genetic polymorphism using the human liver bank as a model system. Drug Metab Rev, 42, 209–24.

Curtis SW, Washburn T, Sewall C, DiAugustine R, Lindzey J, Couse JF, Korach KS. (1996). Physiological coupling of growth factor and steroid receptor signaling pathways: estrogen receptor knockout mice lack estrogen-like response to epidermal growth factor. Proc Natl Acad Sci USA, 93, 12626–30.

Das SK, Taylor JA, Korach KS, Paria BC, Dey SK, Lubahn DB. (1997). Estrogenic responses in estrogen receptor-alpha deficient mice reveal a distinct estrogen signaling pathway. Proc Natl Acad Sci USA, 94, 12786–91.

de Wildt SN, Kearns GL, Leeder JS, van den Anker JN. (1999). Glucuronidation in humans. Pharmacogenetic and developmental aspects. Clin Pharmacokinet, 36, 439–52.

Deakin SP, James RW. (2004). Genetic and environmental factors modulating serum concentrations and activities of the antioxidant enzyme paraoxonase-1. Clin Sci (Lond), 107, 435–47.

Dellarco V, Fenner-Crisp PA. (2012). Mode of action: moving toward a more relevant and efficient assessment paradigm. J. Nutr, 142, 2192S–8S.

Dietrich DR. (2010). Courage for simplification and imperfection in the 21st century assessment of "Endocrine disruption". ALTEX, 27, 264–78.

Dougherty EJ, Guo C, Simons SS Jr, Chow, CC. (2012). Deducing the temporal order of cofactor function in ligand-regulated gene transcription: theory and experimental verification. PLoS One, 7, e30225.

EC (European Commission Joint Research Centre). (2003). Technical Guidance Document on Risk Assessment, Part II.

EC (European Commission) 2008. Regulation (European Commission) No, 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures amending and repealing Directives 67/548/EEC and 1999/45/EEC, and amending Regulation (EC) No. 1907/2006.

Edginton AN, Ritter L. (2009). Predicting plasma concentrations of bisphenol A in children younger than 2 years of age after typical feeding schedules, using a physiologically base toxicokinetic model. Environ Health Perspect, 117, 645–52.

EFSA (European Food Safety Authority). (2005). Opinion of a Scientific Committee on a request from EFSA related to A Harmonised Approach for Risk Assessment of Substances Which Are Both Genotoxic and Carcinogenic.

Elcombe CR, Peffer RC, Wolf DC, Bailey J, Bars R, Bell D, et al. (2014). Mode of action and human relevance analysis for nuclear receptor-

DOI 10.3109/10408444.2014.931925

mediated liver toxicity: a case study with phenobarbital as a model constitutive androstane receptor (CAR) activator. Crit Rev Toxicol, 44, 64–82.

Evans MV, Dowd SM, Kenyon EM, Hughes MF, El-Masri HA. (2008). A physiologically based pharmacokinetic model for intravenous and ingested dimethylarsinic acid in mice. Toxicol Sci, 104, 250–60.

Farley DB, Ford SP, Rosazza, JP. (1992). Increase in uterine peroxidase activity in the rat uterus during oestrogen hyperaemia. J Reprod Fertil, 95, 551–8.

Faustman EM, Ponce RA, Seeley MR, Whittaker SG. (1997). Experimental approaches to evaluate mechanisms of developmental toxicity. Handbook of developmental toxicology. New York: CRC Press, Inc;,

Fisher JW, Twaddle NC, Vanlandingham M, Doerge DR. (2011). Pharmacokinetic modeling: prediction and evaluation of route dependent dosimetry of bisphenol A in monkeys with extrapolation to humans. Toxicol Appl Pharmacol, 257, 122–36.

Furr BJ, Jordan VC. (1984). The pharmacology and clinical uses of tamoxifen. Pharmacol Ther, 25, 127–205.

Gamble MV, Hall MN. (2012). Relationship of creatinine and nutrition with arsenic metabolism. Environ Health Perspect, 120, A145–6.

Garabrant DH, Aylward LL, Berent S, Chen Q, Timchalk C, Burns CJ, et al. (2009). Cholinesterase inhibition in chlorpyrifos workers: characterization of biomarkers of exposure and response in relation to urinary TCPy. J Expo Sci Environ Epidemiol, 19, 634–42.

Geens T, Aerts D, Berthot C, Bourguignon JP, Goeyens L, Lecomte P, et al. (2012). A review of dietary and non-dietary exposure to bisphenol-A. Food Chem Toxicol, 50, 3725–40.

Giangrande PH, Kimbrel EA, Edwards DP, McDonnell DP. (2000). The opposing transcriptional activities of the two isoforms of the human progesterone receptor are due to differential cofactor binding. Mol Cell Biol, 20, 3102–15.

Girard H, Lévesque E, Bellemare J, Journault K, Caillier B, Guillemette C. (2007). Genetic diversity at the UGT1 locus is amplified by a novel 3' alternative splicing mechanism leading to nine additional UGT1A proteins that act as regulators of glucuronidation activity. Pharmacogenet Genomics, 17, 1077–89.

Golden RJ, Noller KL, Titus-Ernstoff L, Kaufman RH, Mittendorf R, Stillman R, Reese EA. (1998). Environmental endocrine modulators and human health: an assessment of the biological evidence. Crit Rev Toxicol, 28, 109–227.

Gori GB. (2013). Regulatory forum opinion piece: long-term animal bioassays: is the end near? Toxicol Pathol, 41, 805–7.

Gorski J, Stormshak F, Harris J, Wertz N. (1977). Hormone regulation of growth: stimulatory and inhibitory influences of estrogens on DNA synthesis. J Toxicol Environ Health, 3, 271–9.

Goutelle S, Maurin M, Rougier F, Barbaut X, Bourguignon L, Ducher M, Maire P. (2008). The Hill equation: a review of its capabilities in pharmacological modelling. Fundam Clin Pharmacol, 22, 633–48.

Guillemette C, Lévesque E, Harvey M, Bellemare J, Menard V. (2010). UGT genomic diversity: beyond gene duplication. Drug Metab Rev, 42, 24–44.

Guzelian PS, Victoroff MS, Halmes NC, James RC, Guzelian CP. (2005). Evidence-based toxicology: a comprehensive frame-work for causation. Hum Exp Toxicol, 24, 161–201.

Hanahan D, Weinberg RA. (2000). The hallmarks of cancer. Cell, 100, 57–70.

Hanahan D, Weinberg RA. (2011). Hallmarks of cancer: the next generation. Cell, 144, 646–74.

Hartung T, McBride M. (2011). Food for thought … on mapping the human toxome. ALTEX, 28, 83–93.

Health Canada. (1999). Voluntary pesticide resistance-management labelling based on target site/mode of action.

Health Canada. (2009). Guidelines for Canadian drinking water quality: guideline technical document—Benzene.

Health Canada. (2011). Guidelines for Canadian drinking water quality: guideline technical document— N-Nitrosodimethylamine (NDMA).

Heneweer M, Houtman R, Poortman J, Groot M, Maliepaard C, Peijnenburg A. (2007). Estrogenic effects in the immature rat uterus after dietary exposure to ethinylestradiol and zearalenone using a systems biology approach. Toxicol Sci, 99, 303–14.

Hengstler JG, Foth H, Gebel T, Kramer PJ, Lilienblum W, Schweinfurth et al. (2011). Critical evaluation of key evidence on the human health hazards of exposure to bisphenol A. Crit Rev Toxicol, 41, 263–91.

Hill AB. (1965). The Environment and disease: association or Causation? . Proc R Soc Med, 58, 295–300.

Hill AV. (1910). The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves. J. Physiol Proceedings: iv-vii.

Hinderliter PM, Price PS, Bartels MJ, Timchalk C, Poet TS. (2011). Development of a source-to-outcome model for dietary exposures to insecticide residues: an example using chlorpyrifos. Regul Toxicol Pharmacol, 61, 82–92.

Huen K, Harley K, Bradman A, Eskenazi B, Holland N. (2010). Longitudinal changes in PON1 enzymatic activities in Mexican-American mothers and children with different genotypes and haplotypes. Toxicol Appl Pharmacol, 244, 181–9.

IOM (Institute of Medicine). (2010). Evaluation of Biomarkers and Surrogate Endpoints in Chronic Disease. The National Academies Press.

John S, Sabo PJ, Thurman RE, Sung MH, Biddie SC, Johnson TA, et al. (2011). Chromatin accessibility pre-determines glucocorticoid receptor binding patterns. Nat Genet, 43, 264–8.

Judson R, Houck K, Martin M, Knudsen T, Thomas RS, Sipes N, et al. (2014). In Vitro and Modelling Approaches to Risk Assessment from the U.S. Environmental Protection Agency ToxCast Programme. Basic Clin Pharmacol Toxicol, 115, 50–76.

Julien E, Boobis AR, Olin SS, Ilsi Research Foundation Threshold Working Group. (2009). The Key Events Dose-Response Framework: a cross-disciplinary mode-of-action based approach to examining dose-response and thresholds. Crit Rev Food Sci Nutr, 49, 682–9.

Kaye AM, Icekson I, Lindner HR. (1971). Stimulation by estrogens of ornithine and S-adenosylmethionine decarboxylases in the immature rat uterus. Biochim Biophys Acta, 252, 150–9.

Kirman CR, Hays SM, Aylward LL, Suh M, Harris MA, Thompson CM, et al. (2012). Physiologically based pharmacokinetic model for rats and mice orally exposed to chromium. Chem Biol Interact, 200, 45–64.

Klotz DM, Hewitt SC, Ciana P, Raviscioni M, Lindzey JK, Foley J, et al. (2002). Requirement of estrogen receptor-alpha in insulin-like growth factor-1 (IGF-1)-induced uterine responses and in vivo evidence for IGF-1/estrogen receptor cross-talk. J Biol Chem, 277, 8531–537.

Kobayashi S, Stice JP, Kazmin D, Wittmann BM, Kimbrel EA, Edwards DP, et al. (2010). Mechanisms of progesterone receptor inhibition of inflammatory responses in cellular models of breast cancer. Mol Endocrinol, 24, 2292–302.

Koshland DE. (1996). The structural basis of negative cooperativity: receptors and enzymes. Curr Opin Struct Biol, 6, 757–61.

Koshland DE, Hamadani K. (2002). Proteomics and models for enzyme cooperativity. J Biol Chem, 277, 46841–4.

Kraus WL, Weis KE, Katzenellenbogen BS. (1995). Inhibitory cross-talk between steroid hormone receptors: differential targeting of estrogen receptor in the repression of its transcriptional activity by agonist- and antagonist-occupied progestin receptors. Mol Cell Biol, 15, 1847–57.

Krekels EH, Danhof M, Tibboel D, Knibbe CA. (2012). Ontogeny of hepatic glucuronidation; methods and results. Curr Drug Metab, 13, 728–43.

Landesmann B, Mennecozzi M, Berggren E, Whelan M. (2013). Adverse outcome pathway-based screening strategies for an animal-free safety assessment of chemicals, Altern Lab Anim, 41, 461–71.

Leary FJ, Resseguie LJ, Kurland LT, O'brien PC, Emslander RF, Noller KL. (1984). Males exposed in utero to diethylstilbestrol. JAMA, 252, 2984–9.

Lehman-McKeeman LD, Rodriguez PA, Takigiku R, Caudill D, Fey ML. (1989). d-Limonene-induced male rat-specific nephrotoxicity: Evaluation of the association between d-limonene and alpha 2u-globulin. Toxicol Appl Pharmacol, 99, 250–9.

Levin E, Actis AM, Lopez S. (1993). Characterization of rat uterine estrogen receptors in vivo. J Steroid Biochem Mol Biol, 44, 277–85.

Levine M. (2011). Paused RNA polymerase II as a developmental checkpoint. Cell, 145, 502–11.

Levitzki A, Koshland DE. (1969). Negative cooperativity in regulatory enzymes. Proc Natl Acad Sci U. S. A, 62, 1121–8.

Liteplo RG, Meek ME. (2003). Inhaled formaldehyde: Exposure estimation, hazard characterization and exposure-response analysis. J Toxicol Environ Health B Crit Rev, 6, 85–114.

Lutz WK. (2001). Susceptibility differences in chemical carcinogenesis lincarize the dose-response relationship: threshold doses can be defined only for individuals. Mutat Res, 482, 71–6.

Lutz WK, Lutz RW. (2009). Statistical model to estimate a threshold dose and its confidence limits for the analysis of sublinear dose-response relationships, exemplified for mutagenicity data. Mutat Res, 678, 118–22.

Lyttle CR, Desombre ER. (1977). Uterine peroxidase as a marker for estrogen action. Proc Natl Acad Sci U. S. A, 74, 3162–6.

Maier A, Kohrman-Vincent M, Hertzberg R, Allen B, Haber LT, Dourson M. (2012). Critical review of dose-response options for F344 rat

mammary tumors for acrylamide - additional insights based on mode of action. Food Chem Toxicol, 50, 1763–75.

Mazur CS, Kenneke JF, Hess-Wilson JK, Lipscomb JC. (2010). Differences between human and rat intestinal and hepatic bisphenol A glucuronidation and the influence of alamethicin on in vitro kinetic measurements. Drug Metab Dispos, 38, 2232–8.

Meek ME. (2008). Recent developments in frameworks to consider human relevance of hypothesized modes of action for tumours in animals. Environ Mol Mutagen, 49, 110–16.

Meek ME, Boobis A, Cote I, Dellarco V, Fotakis G, Munn S, et al. (2014b). New developments in the evolution and application of the WHO/IPCS framework on mode of action/species concordance analysis. J Appl Toxicol, 34, 1–18.

Meek ME, Bucher JR, Cohen SM, Dellarco V, Hill RN, Lehman-Mckeeman LD, et al. (2003). A framework for human relevance analysis of information on carcinogenic modes of action. Crit. Rev Toxicol, 33, 591–653.

Meek ME, Palermo CM, Bachman AN, North CM, Jeffrey Lewis R. (2014a). Mode of action human relevance (species concordance) framework: Evolution of the Bradford Hill considerations and comparative analysis of weight of evidence. J Appl Toxicol, 34, 595–606.

Mercke Odeberg J, Andradem J, Holmberg K, Hoglund P, Malmqvist U, Odeberg J. (2006). UGT1A polymorphisms in a Swedish cohort and a human diversity panel, and the relation to bilirubin plasma levels in males and females. Eur J Clin Pharmacol, 62, 829–37.

Mielke H, Gundert-Remy U. (2009). Bisphenol A levels in blood depend on age and exposure. Toxicol Lett, 190, 32–40.

Mileson BE, Chambers JE, Chen WL, Dettbarn W, Ehrich M, Eldefrawi AT, et al. (1998). Common mechanism of toxicity: a case study of organophosphorus pesticides. Toxicol Sci, 41, 8–20.

Miyagi SJ, Collier AC. (2011). The development of UDP-glucuronosyltransferases 1A1 and 1A6 in the pediatric liver. Drug Metab Dispos, 39, 912–19.

Murrell JA, Portier CJ, Morris RW. (1998). Characterizing dose-response: I: Critical assessment of the benchmark dose concept. Risk Anal, 18, 13–26.

Nachman RA, Fox SD, Golden C, Sibinga E, Veenstra TD, Groopman JD, Lees PSJ. (2013). Urinary Free Bisphenol A and Bisphenol A-Glururonide Concentrations in Newborns. J Pediatr, 162, 870–2.

Naciff JM, Overmann GJ, Torontali SM, Carr GJ, Tiesman JP, Richardson BD, Daston GP. (2003). Gene expression profile induced by 17 alpha-ethynyl estradiol in the prepubertal female reproductive system of the rat. Toxicol Sci, 72, 314–30.

Navarro SL, Chen Y, Li L, Li SS, Chang JL, Schwarz Y, et al. (2011). UGT1A6 and UGT2B15 polymorphisms and acetaminophen conjugation in response to a randomized, controlled diet of select fruits and vegetables. Drug Metab Dispos, 39, 1650–7.

Navarro SL, Peterson S, Chen C, Makar KW, Schwarz Y, King IB, Li SS, Li L, Kestin M, Lampe JW. (2009). Cruciferous vegetable feeding alters UGT1A1 activity: diet- and genotype-dependent changes in serum bilirubin in a controlled feeding trial. Cancer Prev Res (Phila. Pa), 2, 345–52.

Notides AC, Lerner N, Hamilton DE. (1981). Positive cooperativity of the estrogen receptor. Proc Natl Acad Sci U. S. A, 78, 4926–30.

NRC (National Research Council). (1983). Risk Assessment in the Federal Government: Managing the Process.

NRC (National Research Council). (2007). Toxicity Testing in the Twenty-First Century.

NRC (National Research Council) 2009. Science and Decisions: Advancing Risk Assessment. Washington, DC: The National Academies Press.

NRC (National Research Council). (2011). Review of the Environmental Protection Agency's Draft IRIS Assessment of Formaldehyde.

NTP (National Toxicology Program). (2006). NTP technical report on the toxicology and carcinogenesis studies of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) (CAS No 1746-01-6) in female Harlan Sprague-Dawley rats (Gavage Studies). Nat. Toxicol. Program. Tech. Rep. Ser TR, 521.

OECD (Organisation for Economic Co-Operation and Development). (2003). Detailed Background Review of the Uterotrophic Bioassay. ENV/JM/MONO 2003.

OECD (Organisation for Economic Co-Operation and Development). (2013). Guidance Document on Developing and Assessing Adverse Outcome Pathways.

Ong KM, Blackford JA, Kagan BL, Simons SS Jr., Chow CC. (2010). A theoretical framework for gene induction and experimental comparisons. Proc Natl Acad Sci U. S. A, 107, 7107–12.

Parisi F, Sonderegger B, Wirapati P, Delorenzi M, Naef F. (2009). Relationship between estrogen receptor alpha location and gene induction reveals the importance of downstream sites and cofactors. BMC Genomics, 10, 381.

Patlewicz G, Simon T, Goyak K, Phillips RD, Rowlands JC, Seidel S, Becker RA. (2013). Use and validation of HT/HC assays to support 21st century toxicity evaluations. Regul Toxicol Pharmacol, 65, 259–68,

Phillips CV, Goodman KJ. (2004). The missed lessons of Sir Austin Bradford Hill. Epidemiol Perspect Innov, 1, 1–5.

Phillips CV, Goodman KJ. (2006). Causal criteria and counterfactuals; nothing more (or less) than scientific common sense. Emerg Themes Epidemiol, 3, 1–7.

Piccirillo VJ, Bird MG, Lewis RJ, Bover WJ. (2012). Preliminary evaluation of the human relevance of respiratory tumors observed in rodents exposed to naphthalene. Regul Toxicol Pharmacol, 62, 433–40.

Popper K. (1959). In The Logic of Scientific Discovery. New York, NY: Republished by Ruteledge,

Price PS, Schnelle KD, Cleveland CB, Bartels MJ, Hinderliter PM, Timchalk C, Poet TS. (2011). Application of a source-to-outcome model for the assessment of health impacts from dietary exposures to insecticide residues. Regul Toxicol Pharmacol, 61, 23–31.

Rao VS, Chaves MC, Ribeiro RA. (1995). Nitric oxide synthase inhibition and the uterotrophic response to oestrogen in immature rats. J Reprod Fertil, 105, 303–6.

Reiss R, Neal B, Lamb JC, Juberg DR. (2012). Acetylcholinesterase inhibition dose-response modeling for chlorpyrifos and chlorpyrifos-oxon. Regul Toxicol Pharmacol, 63, 124–31.

Rhomberg LR, Bailey LA, Goodman JE. (2010). Hypothesis-based weight of evidence: a tool for evaluating and communicating uncertainties and inconsistencies in the large body of evidence in proposing a carcinogenic mode of action–naphthalene as an example. Crit Rev Toxicol, 40, 671–96.

Richter RJ, Jarvik GP, Furlong CE. (2009). Paraoxonase 1 (PON1) status and substrate hydrolysis. Toxicol Appl Pharmacol, 235, 1–9.

Rock W, Rosenblat M, Miller-Lotan R, Levy AP, Elias M, Aviram M. (2008). Consumption of wonderful variety pomegranate juice and extract by diabetic patients increases paraoxonase 1 association with high-density lipoprotein and stimulates its catalytic activities. J Agric Food Chem, 56, 8704–13.

Rodgers IS, Baetcke KP. (1993). Interpretation of male rat renal tubule tumors. Environ Health Perspect, 101, 45–52.

Rosenfeld CR, Cox BE, Roy T, Magness RR. (1996). Nitric oxide contributes to estrogen-induced vasodilation of the ovine uterine circulation. J Clin Invest, 98, 2158–66.

Ross AC, Russell RM, Miller SA, Munro IC, Rodricks JV, Yetley EA, Julien E. (2009). Application of a key events dose-response analysis to nutrients: a case study with vitamin A (retinol). Crit Rev Food Sci Nutr, 49, 708–17.

Rowlands JC, Sander M, Bus JS, Futuretox Organizing Committee. (2014). FutureTox: building the road for 21st century toxicology and risk assessment practices. Toxicol Sci, 137, 269–77.

Sand S, Von Rosen D, Victorin K, Filipsson AF. (2006). Identification of a critical dose level for risk assessment: developments in benchmark dose analysis of continuous endpoints. Toxicol Sci, 90, 241–51.

Saracino MR, Bigler J, Schwarz Y, Chang JL, Li S, Li L, et al. (2009). Citrus fruit intake is associated with lower serum bilirubin concentration among women with the UGT1A1*28 polymorphism. J Nutr, 139, 555–60.

Sawilowsky S. (2002). Fermat, Schubert, Einstein and Behrens-Fisher: The Probable Difference Between Two Means When $\varsigma 12 \neq \varsigma 22$. J Modern Appl Stat Meth, 1, 461–72.

Seed J, Carney EW, Corley RA, Crofton KM, Desesso JM, Foster PM, et al. (2005). Overview: Using mode of action and life stage information to evaluate the human relevance of animal toxicity data. Crit Rev Toxicol, 35, 664–72.

Sen B, Wang A, Hester SD, Robertson JL, Wolf DC. (2005). Gene expression profiling of responses to dimethylarsinic in female F344 rat urothelium. Toxicology, 215, 214–26.

Seok J, Warren HS, Cuenca AG, Mindrinos MN, Baker HV, Xu W, et al. (2013). Genomic responses in mouse models poorly mimic human inflammatory diseases. Proc Natl Acad Sci U.S.A, 110, 3507–12.

Sierksma A, Van Der Gaag MS, Van Tol A, James RW, Hendriks HF. (2002). Kinetics of HDL cholesterol and paraoxonase activity in moderate alcohol consumers. Alcohol Clin Exp Res, 26, 1430–5.

Silkworth JB, Koganti A, Illouz K, Possolo A, Zhao M, Hamilton SB. (2005). Comparison of TCDD and PCB CYP1A induction sensitivities in fresh hepatocytes from human donors, sprague-dawley rats, and rhesus monkeys and HepG2 cells. Toxicol Sci, 87, 508–19.

DOI 10.3109/10408444.2014.931925

Simon T, Aylward LL, Kirman CR, Rowlands JC, Budinsky RA. (2009). Estimates of cancer potency of 2,3,7,8-tetrachlorodibenzo(p)dioxin using linear and nonlinear dose-response modeling and toxicokinetics. Toxicol Sci, 112, 490–506.

Simons SS Jr. (2008). What goes on behind closed doors: physiological versus pharmacological steroid hormone actions. Bioessays, 30, 744–56.

Simons SS Jr. (2010). Glucocorticoid receptor cofactors as therapeutic targets. Curr Opin Pharmacol, 10, 613–19.

Simons SS Jr, Chow CC. (2012). The road less traveled: new views of steroid receptor action from the path of dose-response curves. Mol Cell Endocrinol, 348, 373–82.

Slob W, Setzer RW. (2014). Shape and steepness of toxicological dose-response relationships of continuous endpoints. Crit Rev Toxicol, 44, 270–97.

Smith JN, Timchalk C, Bartels MJ, Poet TS. (2011). In vitro age-dependent enzymatic metabolism of chlorpyrifos and chlorpyrifos-oxon in human hepatic microsomes and chlorpyrifos-oxon in plasma. Drug. Metab. Dispos, 39, 1353–62.

Sonich-Mullin C, Fielder R, Wiltse J, Baetcke K, Dempsey J, Fenner-Crisp P, et al. (2001). IPCS conceptual framework for evaluating a mode of action for chemical carcinogenesis. Regul Toxicol Pharmacol, 34, 146–52.

Stice SL, Ford SP, Rosazza JP, Van Orden DE. (1987a). Interaction of 4-hydroxylated estradiol and potential-sensitive Ca2 + channels in altering uterine blood flow during the estrous cycle and early pregnancy in gilts. Biol Reprod, 36, 369–75.

Stice SL, Ford SP, Rosazza JP, Van Orden DE. (1987b). Role of 4-hydroxylated estradiol in reducing Ca2 + uptake of uterine arterial smooth muscle cells through potential-sensitive channels. Biol Reprod, 36, 361–8.

Strassburg CP, Oldhafer K, Manns MP, Tukey RH. (1997). Differential expression of the UGT1A locus in human liver, biliary, and gastric tissue: identification of UGT1A7 and UGT1A10 transcripts in extrahepatic tissue. Mol Pharmacol, 52, 212–20.

Sun Y, Tao YG, Kagan BL, He Y, Simons Jr, SS. (2008). Modulation of transcription parameters in glucocorticoid receptor-mediated repression. Mol Cell Endocrinol, 295, 59–69.

Susser M. (1986). The logic of Sir Karl Popper and the practice of epidemiology. Am. J. Epidemiol, 124, 711–18.

Taylor SL, Gendel SM, Houben GF, Julien E. (2009). The Key Events Dose-Response Framework: a foundation for examining variability in elicitation thresholds for food allergens. Crit Rev Food Sci Nutr, 49, 729–39.

Teeguarden JG, Wacchter JM, Clewell HJ, Covington TR, Barton HA. (2005). Evaluation of oral and intravenous route pharmacokinetics, plasma protein binding, and uterine tissue dose metrics of bisphenol A: a physiologically based pharmacokinetic approach. Toxicol Sci, 85, 823–38.

Thomas DJ. (2007). Molecular processes in cellular arsenic metabolism. Toxicol Appl Pharmacol, 222, 365–73.

Thomas RS, Black M, Li L, Healy E, Chu TM, Bao W, et al. (2012). A Comprehensive Statistical Analysis of Predicting In Vivo Hazard Using High-Throughput In Vitro Screening. Toxicol Sci, 128, 398–417.

Thomas RS, Philbert MA, Auerbach SS, Wetmore BA, Devito MJ, Cote I, et al. (2013). Incorporating new technologies into toxicity testing and risk assessment: moving from 21st century vision to a data-driven framework. Toxicol Sci, 136, 4–18.

Thompson CM, Kirman CR, Proctor DM, Haws LC, Suh M, Hays SM, et al. (2014). A chronic oral reference dose for hexavalent chromium-induced intestinal cancer. J Appl Toxicol, 34, 525–36.

Timchalk C, Kousba A, Poet TS. (2002a). Monte Carlo analysis of the human chlorpyrifos-oxonase (PON1) polymorphism using a physiologically based pharmacokinetic and pharmacodynamic (PBPK/PD) model. Toxicol Lett, 135, 51–9.

Timchalk C, Nolan RJ, Mendrala AL, Dittenber DA, Brzak KA, Mattsson JL. (2002b). A Physiologically based pharmacokinetic and pharmacodynamic (PBPK/PD) model for the organophosphate insecticide chlorpyrifos in rats and humans. Toxicol Sci, 66, 34–53.

Tukey JW, Ciminera JL, Heyse JF. (1985). Testing the Statistical Certainty of a Response to Increasing Doses of a Drug. Biometrics, 41, 295–301.

USEPA (United States Environmental Protection Agency). (2005a). Guidelines for Carcinogen Risk Assessment.

USEPA (United States Environmental Protection Agency). (2005b). Supplemental Guidance for Assessing Susceptibility from Early-Life Exposure to Carcinogens.

USEPA (United States Environmental Protection Agency). (2005c). Science Issue Paper: Mode of Carcinogenic Action for Cacodylic Acid V (Dimethylarsinic Acid, DMA) and Recommendations for Dose Response Extrapolation.

USEPA (United States Environmental Protection Agency). (2006). Revised Reregistration Eligibility Decision for MSMA, DSMA, CAMA, and Cacodylic Acid.

USEPA (United States Environmental Protection Agency). (2007). Framework for Determining a Mutagenic Mode of Action for Carcinogenicity: Using EPA's 2005 Cancer Guidelines and Supplemental Guidance for Assessing Susceptibility to Early-Life Exposure to Carcinogens.

USEPA (United States Environmental Protection Agency). (2011). DRAFT Guidance for Applying Quantitative Data to Develop Data-Derived Extrapolation Factors for Interspecies and Intraspecies Extrapolation.

van Gemert M, Dourson M, Moretto A, Watson M. (2001). Use of human data for the derivation of a reference dose for chlorpyrifos. Regul. Toxicol. Pharmacol, 33, 110–16.

Vink R, Mikkers J, Bouwman T, Marquart H, Kroese D. (2010). Use of read-across and tiered exposure assessment in risk assessment under REACH – A case study on a phase-in substance. Regul Toxicol Pharmacol, 58, 64–71.

Vinken M. (2013). The adverse outcome pathway concept: a pragmatic tool in toxicology. Toxicology, 312, 158–65.

Waddell WJ. (2005). Comparisons of thresholds for carcinogenicity on linear and logarithmic dosage scales. Hum Exp Toxicol, 24, 325–32.

Waddell WJ. (2008). Thermodynamic basis for expressing dose logarithmically. Toxicol Appl Pharmacol, 2, 156–7.

Wagner JG. (1968). Kinetics of pharmacologic response. I. Proposed relationships between response and drug concentration in the intact animal and man. J Theor Biol, 20, 173–201.

Wen DX, Xu YF, Mais DE, Goldman ME, Mcdonnell DP. (1994). The A and B isoforms of the human progesterone receptor operate through distinct signaling pathways within target cells. Mol. Cell. Biol, 14, 8356–64.

Willhite CC, Ball GL, Mclellan CJ. (2008). Derivation of a bisphenol A oral reference dose (RfD) and drinking-water equivalent concentration. J Toxicol Environ Health B Crit Rev, 11, 69–146.

Winuthayanon W, Hewitt SC, Orvis GD, Behringer RR, Korach KS. (2010). Uterine epithelial estrogen receptor α is dispensable for proliferation but essential for complete biological and biochemical responses. Proc Natl Acad Sci USA, 107, 19272–7.

World Health Organization – International Programme on Chemical Safety (WHO-IPCS). (2005). Chemical-Specific Adjustment Factors for Interspecies Differences and Human Variability: Guidance Document for Use of Data in Dose/Concentration Response Assessment.

Yamasaki K, Noda S, Imatanaka N, Yakabe Y. (2004). Comparative study of the uterotrophic potency of 14 chemicals in a uterotrophic assay and their receptor-binding affinity. Toxicol Lett, 146, 111–20.

Yamasaki K, Takeyoshi M, Yakabe Y, Sawaki M, Imatanaka N, Takatsuki M. (2002). Comparison of reporter gene assay and immature rat uterotrophic assay of twenty-three chemicals. Toxicology, 170, 21–30.

Yamasaki K, Takeyoshi M, Yakabe Y, Sawaki M, Takatsuki M. (2003). Comparison of the reporter gene assay for ER-alpha antagonists with the immature rat uterotrophic assay of 10 chemicals. Toxicol Lett, 142, 119–31.

Ye X, Zhou X, Wong LY, Calafat AM. (2012). Concentrations of bisphenol a and seven other phenols in pooled sera from 3–11 year old children: 2001–2002 national health and nutrition examination survey. Environ Sci Technol, 46, 12664–71.

Zhang Z, Sun Y, Cho YW, Chow CC, Simons SS. (2012). PA1: a New Competitive Decelerator Acting at More Than One Step To Impede Glucocorticoid Receptor-Mediated Transactivation. J. Biol. Chem, 288, 42–58.

Zhang Z, Teng CT. (2001). Estrogen receptor alpha and estrogen receptor-related receptor alpha1 compete for binding and coactivator. Mol Cell Endocrinol, 172223–33.

Supplementary material available online

Supplementary Content to be found online at http://informahealthcare.com/doi/abs/10.3109/10408444.2014.931925.

Review Article

Journal of Applied Toxicology

Received: 18 September 2013, Accepted: 19 September 2013 Published online in Wiley Online Library: 25 October 2013

(wileyonlinelibrary.com) DOI 10.1002/jat.2949

New developments in the evolution and application of the WHO/IPCS framework on mode of action/species concordance analysis†

M. E. Meek[a], A. Boobis[b], I. Cote[c], V. Dellarco[d], G. Fotakis[e], S. Munn[f], J. Seed[g] and C. Vickers[h]*

ABSTRACT: The World Health Organization/International Programme on Chemical Safety mode of action/human relevance framework has been updated to reflect the experience acquired in its application and extend its utility to emerging areas in toxicity testing and non-testing methods. The underlying principles have not changed, but the framework's scope has been extended to enable integration of information at different levels of biological organization and reflect evolving experience in a much broader range of potential applications. Mode of action/species concordance analysis can also inform hypothesis-based data generation and research priorities in support of risk assessment. The modified framework is incorporated within a roadmap, with feedback loops encouraging continuous refinement of fit-for-purpose testing strategies and risk assessment. Important in this construct is consideration of dose–response relationships and species concordance analysis in weight of evidence. The modified Bradford Hill considerations have been updated and additionally articulated to reflect increasing experience in application for cases where the toxicological outcome of chemical exposure is known. The modified framework can be used as originally intended, where the toxicological effects of chemical exposure are known, or in hypothesizing effects resulting from chemical exposure, using information on putative key events in established modes of action from appropriate *in vitro* or *in silico* systems and other lines of evidence. This modified mode of action framework and accompanying roadmap and case examples are expected to contribute to improving transparency in explicitly addressing weight of evidence considerations in mode of action/species concordance analysis based on both conventional data sources and evolving methods. Copyright © 2013 John Wiley & Sons, Ltd. The World Health Organization retains copyright and all other rights in the manuscript of this article as submitted for publication.

Keywords: key events; mode of action; adverse outcome pathway; human relevance framework; modified Bradford Hill considerations; weight of evidence approach; species concordance analysis; cellular response; tissue response; molecular target

Introduction

The mode of action/human relevance framework was developed in initiatives of the International Programme on Chemical Safety (IPCS) of the World Health Organization (WHO) (Boobis *et al.*, 2006, 2008; Sonich-Mullin *et al.*, 2001) and the International Life Sciences Institute Risk Sciences Institute (ILSI-RSI) (Meek *et al.*, 2003; Seed *et al.*, 2005). It derives from earlier work on mode of action in animals by the US Environmental Protection Agency (US EPA, 1996, 2005a) and has involved large numbers of scientists internationally.

Previous development of the mode of action/human relevance framework is described in the publications mentioned above and summarized more recently in Meek and Klaunig (2010). The framework has been illustrated by an increasing number of case studies (more than 30 currently) demonstrating the value of mode of action in evaluating human relevance and life stage susceptibility and guiding dose–response assessment. Documented examples are presented in Table 1. The contribution of the framework has been recognized by the Society of Toxicology, and the framework has been adopted by several international and national organizations and agencies to increase transparency in the assessment of weight of evidence and identification of critical data needs (Meek, 2008, 2009; Meek *et al.*, 2008).

The framework continues to evolve as experience increases in its application to consider systematically the weight of evidence

*Correspondence to: C. Vickers, International Programme on Chemical Safety, World Health Organization, 20 Avenue Appia, CH-1211 Geneva 27, Switzerland. Email: vickersc@who.int

† This publication contains the collective views of an international group of experts and does not necessarily represent the decisions or the stated policy of the World Health Organization or the authors' affiliated organizations.

[a]Chemical Risk Assessment, McLaughlin Centre for Population Health Risk Assessment, 1 Stewart Street, Ottawa, Ontario, Canada, K1N 6N5

[b]Department of Medicine, Imperial College London, Hammersmith Campus, Ducane Road, London, W12 0NN, UK

[c]National Center for Environmental Assessment, US Environmental Protection Agency, 1200 Pennsylvania Avenue, NW, Washington DC, 20460, USA

[d]Formerly Office of Pesticide Programs, US Environmental Protection Agency, 1200 Pennsylvania Avenue, NW, Washington DC, 20460, USA

[e]European Chemicals Agency, Annankatu 18, 00120, Helsinki, Finland

[f]Systems Toxicology Unit, Institute for Health and Consumer Protection, Joint Research Centre, European Commission, Via E. Fermi, 2749, I-21027, Ispra, VA, Italy

[g]Office of Pollution Prevention and Toxics, US Environmental Protection Agency, 1200 Pennsylvania Avenue, NW, Washington DC, 20460, USA

[h]International Programme on Chemical Safety, World Health Organization, 20 Avenue Appia, CH-1211, Geneva 27, Switzerland

Table 1. Case studies illustrating various modes of action and implications for dose–response assessment

Mode of action	Case study	Reference
Tumors of various organs associated with mutagenic modes of action	Ethylene oxide	Meek *et al.* (2003)
	4-Aminobiphenyl	Cohen *et al.* (2006a)
Mammary tumors associated with suppression of luteinizing hormone	Atrazine	Meek *et al.* (2003)
Thyroid tumors associated with increased clearance of thyroxine	Phenobarbital	Meek *et al.* (2003)
	Thiazopyr	Dellarco *et al.* (2006)
Bladder tumors associated with the formation of urinary tract calculi	Melamine	Meek *et al.* (2003)
Liver/kidney tumors associated with sustained cytotoxicity and regenerative proliferation	Chloroform	Meek *et al.* (2003)
Acute renal toxicity associated with precipitation of oxalate	Ethylene glycol	Seed *et al.* (2005)
Androgen receptor antagonism and developmental effects	Vinclozolin	Seed *et al.* (2005)
Nasal tumors associated with DNA reactivity and cytotoxicity	Formaldehyde	McGregor *et al.* (2006)

from traditional and evolving methods for assessing toxicity. This includes explicit consideration of the comparative weight of evidence and associated uncertainties for several options for hypothesized modes of action early and throughout the analysis. The critical relevance of the kinetic and dynamic information considered in the mode of action analysis for subsequent characterization of dose–response relationships for effects considered relevant to humans (Boobis *et al.*, 2009; Julien *et al.*, 2009), including choice of chemical-specific adjustment factors (Boobis *et al.*, 2008), has also been amplified. Experience in mode of action analysis has also been instructive in contextualizing appropriate application of information from evolving methods of toxicity testing at different levels of biological organization as a basis for more efficient testing strategies.

Objectives

This paper has been prepared as an addendum to the previous WHO/IPCS guidance on mode of action/human relevance analysis (Boobis *et al.*, 2006, 2008). While the underlying principles and methodology are similar, the guidance has been updated to reflect recent developments. Some of these developments result from advances in toxicity testing and non-testing methods, and some reflect evolving experience in mode of action/species concordance analysis (additionally referred to herein as mode of action analysis). More detailed information on the nature of systematic hypothesis generation and weight of evidence considerations in mode of action analysis with illustrative case examples is included in the earlier publications referenced in Table 1.

This paper also expands the scope of previous manuscripts to reflect increased understanding of the role of mode of action/species concordance analysis in integrating information from different levels of biological organization. In addition, while early focus of mode of action analysis related to increasing transparency in documenting an operative mode of action with a reasonably high degree of confidence as a basis for risk assessment and regulatory decision-making, the current paper addresses a much broader range of contexts. These include implications for priority setting and testing strategies for both individual chemicals and chemical categories where a less refined analysis and/or higher uncertainty may be acceptable. Summaries of cases selected to illustrate examples of broad application in a research/regulatory context are included here. Readers are referred to the cited documentation for more detailed information on the data analysis for these cases.

Both cancer and non-cancer effects are addressed, in recognition that their separation in earlier publications reflected principally evolving experience in mode of action/human relevance analysis rather than variation in conceptual premise. In fact, mode of action analysis facilitates harmonization of cancer and non-cancer assessment. Harmonization in this context refers to a biologically consistent approach to risk assessment for all endpoints, for which exploration of biological linkages is critical to ensuring maximal utility of relevant information. Often, for example, cytotoxicity in an organ is a critical key event that may lead to an increase in cell proliferation and tumors at the same site.

Background/Terminology

Mode of action, as previously defined, is a biologically plausible series of key events leading to an effect (Sonich-Mullin *et al.*, 2001). Originally, mode of action was considered principally in the context of late-stage key cellular, biochemical and tissue events. A key event is an empirically observable step or its marker, which is a necessary element of the mode of action critical to the outcome (i.e., necessary, but not necessarily sufficient in its own right); key events are measurable and reproducible. The mode of action framework is based, then, on the premise that any human health effect caused by exposure to an exogenous substance can be described by a series of causally linked biochemical or biological key events that result in a pathological or other disease outcome. (The term mode of action implies no judgment about adversity of effect, though for risk assessment application, the relevant identified or presumed effects are most often considered adverse.) While originally and often simply conceptualized and illustrated as a linear series of key events, in reality, mode of action involves interdependent networks of events with feedback loops. Disease outcomes are initiated or modified within these networks. Differences in networks between and within human and animal populations account, in part, for interspecies differences and human variability.

Early key events in hypothesized modes of action are most often related to chemical characteristics, i.e., those characteristics of structure and/or physicochemical properties that promote interaction of the substance with biological targets. Later key events are less chemical specific and more often an expected consequence of progression of earlier key events (e.g., regenerative proliferation resulting from cytotoxicity).

Journal of AppliedToxicology

An adverse outcome pathway is conceptually similar to a mode of action. It was initially described by the computational ecotoxicology community (Ankley *et al.*, 2010) and has been adopted within an international initiative to document, develop and assess the completeness of potentially predictive tools for adverse ecological and human health effects (OECD, 2012). A focus of adverse outcome pathways is on the initial associated chemically mediated "molecular initiating event," equivalent to an early key event in a mode of action.

The terms mode of action and adverse outcome pathway should be interchangeable, representing essentially the subdivision of the pathway between exposure and effect in either individuals or populations into a series of hypothesized key events at different levels of biological organization (e.g., molecular, subcellular, cellular, tissue) (Fig. 1). (The term toxicity pathway, introduced by the US National Research Council in 2007 [NRC, 2007], essentially focuses on a subset of early events leading to an effect at the molecular and cellular levels. These events can be considered critical upstream elements of a more expansive mode of action description of how a chemical can affect human health.) The distinction between mode of action and adverse outcome pathway is artificial, a result principally of experience in the human health versus ecological communities, though it has sometimes been stated incorrectly that, unlike adverse outcome pathway, mode of action does not extend from the individual to the population level. It should be noted, though, that the term mode of action, *per se*, does not imply adversity of outcome. Mode of action, as defined here, could apply equally well to effects that are not adverse, such as therapeutic interventions or health benefits (e.g., from nutritional supplements). Also, focus on human health risk assessment has traditionally been on (often later) key events that provide quantitative information relevant to intraspecies and interspecies extrapolation and life stage susceptibility for dose–response analysis, compared with the molecular initiating event in ecological health assessment. For this reason, considerations relevant to weight of evidence analysis may differ.

Appropriately, given their conceptual similarity, it has been proposed that the weight of evidence for both hypothesized modes of action and adverse outcome pathways should draw upon modified Bradford Hill considerations (Hill, 1965). This proposal was based on a desire to increase transparency and consistency in organizing, linking and integrating information at different levels of biological organization into a more efficient, hypothesis-driven approach to chemical data generation and assessment and use of non-test (e.g., read-across and grouping of chemicals) and *in vitro* methods.

However, there are a number of limitations that remain to be addressed in the proposed reliance on modified Bradford Hill considerations for documentation of mode of action where focus has been on the molecular initiating event (i.e., structure–activity modeling). For example, weight of evidence for hypothesized modes of action in human health risk assessment has traditionally relied heavily on the modified Bradford Hill considerations of concordance of dose–response relationships between key and end events. In addition, influential in mode of action analysis is specificity, which in this context has related to experimental verification that a key event is causal. And while experience in mode of action analyses for documented (adverse) effects in human health risk assessment can inform consideration of weight of evidence for hypothesized modes of action or adverse outcome pathways, based on early key or molecular initiating events, to date, information on dose–response concordance and specificity has not been available in characterizing weight of evidence for hypothesized adverse outcome pathways. This detracts considerably from transparency in documentation of their supporting evidence.

Mode of Action/Adverse Outcome Pathways—Levels of Biological Organization

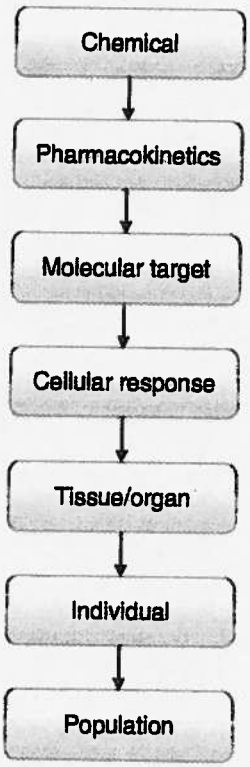


Figure 1. Different levels of biological organization in mode of action analysis. Confidence in an hypothesized mode of action generally increases with increasing evidence at higher levels of biological organization.

Mode of Action Roadmap

There is growing recognition of the need for more efficient methods and strategies to assess the hazards, exposures and risks of the wide array of chemicals to which humans are exposed. This has been reflected in, among others, progressive regulatory mandates in Canada, the European Union and, more recently, the Asian Pacific region to systematically consider priorities for risk management from among all existing chemicals (see, for example, Council of Labor Affairs, Taiwan, 2012; Dellarco *et al.*, 2010; European Commission, 2006; Hughes *et al.*, 2009; Lowell Center for Sustainable Production, 2012; Meek and Armstrong, 2007). This necessitates focus on efficiently prioritized chemicals and endpoints, rather than the traditional time- and resource-intensive series of standard *in vivo* toxicology studies. It also requires the development and integration of information on key events within (hypothesized) modes of action very early in the evaluation process that will enable effective use of data collected from lower levels of biological organization and non-test methods, such as (quantitative) structure–activity relationships ((Q)SAR) and read-across *in vitro* assays.

Figure 2 presents a "mode of action roadmap" to illustrate the iterative process whereby principles and concepts of mode of action analysis can be applied throughout human health risk assessment, with the extent of the analysis being tailored to the issue under consideration. Critical to this more tailored consideration of appropriate testing and assessment strategies is formal, transparent consultation with risk managers, with public accountability, where possible, for the relevant extent of resource investment to address the problem at hand (i.e., problem formulation).

Problem formulation (Fig. 3), the first step in the roadmap (Fig. 2), involves consideration of the risk management scope and goals in relation to relevant exposure scenarios, available resources, urgency of the assessment and the level of uncertainty that is acceptable. This includes consideration of appropriate methods and endpoints for hazard assessment and a mode of action analysis plan tailored to the nature of the decision to be made. For example, decisions concerning chemical prioritization for testing and/or assessment will likely allow for higher levels of uncertainty than those related to establishing regulatory standards. In problem formulation, then, the complexity of the envisaged mode of action analysis is tailored to the context of decision-making; approaches are necessarily flexible and iterative, permitting efficient identification and generation of the essential information to serve as a basis to assess and manage risks appropriately.

The second step in the roadmap (Fig. 2) is to assimilate and consider, in iterative fashion, information on mode of action in the "Modified framework" (see below). This entails hypothesis-based analysis of the weight of evidence for operative key events based on the modified Bradford Hill considerations and qualitative and quantitative concordance of the key events within and between species (Boobis *et al.*, 2006, 2008; Meek *et al.*, 2003; Seed *et al.*, 2005). Early consideration of hypothesis-based key events in the mode of action during problem formulation facilitates incorporation of data from different sources and provides a framework by which it can be organized, integrated and linked at different levels of biological organization (Fig. 3). This includes information generated by evolving methods, such as those targeting cell signaling pathways. The

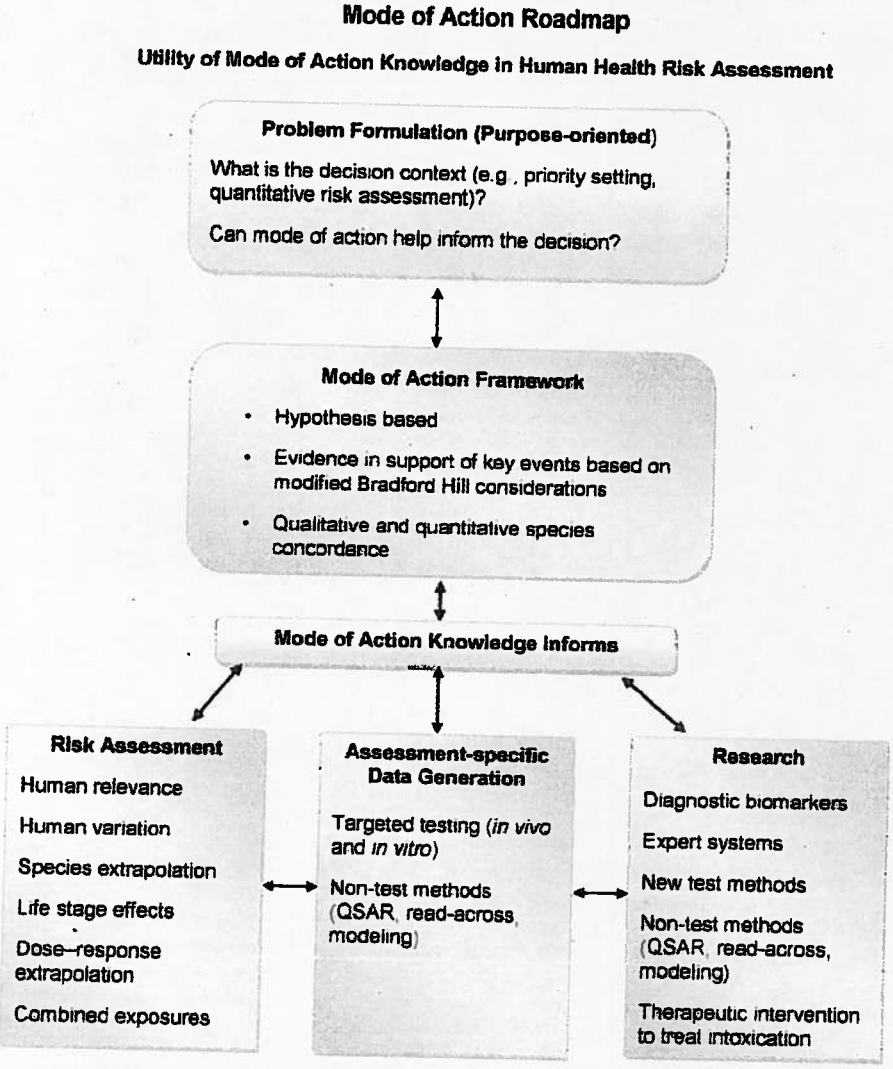


Figure 2. Mode of action roadmap illustrating the use of mode of action knowledge in human health risk assessment. The extent of analysis is tailored to the issue under consideration through iterative analysis and consultation among the assessment, management and research communities.

Journal of Applied Toxicology

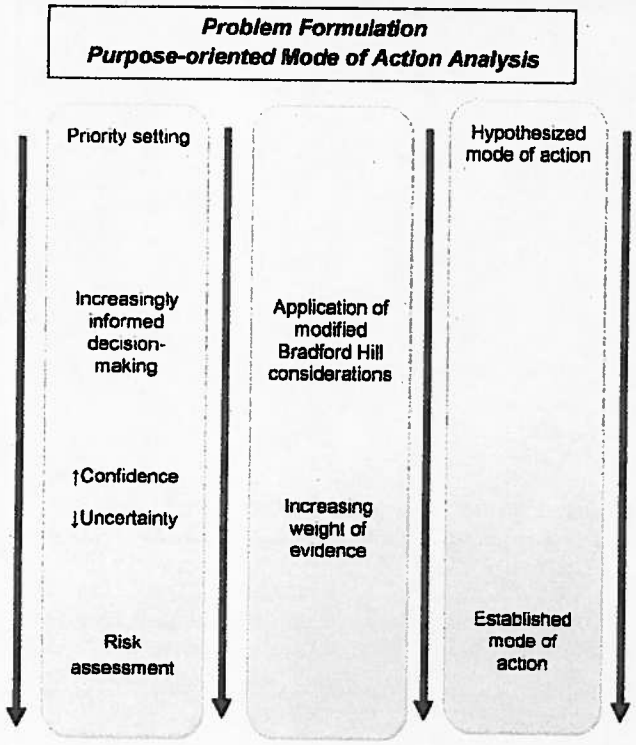


Figure 3. Confidence/uncertainty in "fit for purpose" mode of action/species concordance analysis: correlation of confidence/uncertainty with extent of weight of evidence.

amount of detail and "linearity" characterizing the key events within a hypothesized mode of action can vary as a function of the toxicity of interest, existing knowledge and risk assessment or testing needs.

The mode of action analysis, completed to address the goals outlined during problem formulation, informs one or more of three analytical domains (shown at the bottom of Fig. 2):

(1) risk assessment, including qualitative and quantitative human relevance and variability (e.g., effects at various life stages and within susceptible subgroups), dose–response extrapolation and potential for combined effects of chemicals;
(2) hypothesis-based targeted testing or application of non-test methods to meet the objectives specified in problem formulation, including efficient grouping of chemicals and consideration of read-across, (Q)SAR modeling or appropriate testing within a category approach to fill data needs; and
(3) research priorities relevant to the development of new test and non-test methods, biomarkers and expert systems that feed back to the risk assessment and therapeutic intervention strategies (for intoxication).

As depicted in the roadmap (Fig. 2), mode of action analysis is envisioned as an iterative hypothesis generating and testing process that defines how to assess or test strategically based on risk management needs. As analyses are completed, the problem formulation, testing strategy and risk assessment can be further refined for the decision context.

This iterative process can be illustrated with the following hypothetical example, for which there are considerable data on hazard. While this example draws on a relatively extensive data set, it provides a model for considering significantly fewer data on similar compounds, if they are taken into account from the outset in problem formulation. Initially, a risk manager requests that a risk assessment for the general population be conducted for chemical X, for which exposures of potential concern are those through drinking water. In relatively extensive (traditional) toxicity studies (including a cancer bioassay), chemical X has caused liver tumors in rodents. There is controversy regarding the relevance of this particular tumor type for human health risk assessment, and, based on the preliminary mode of action/species concordance analysis in problem formulation, the risk manager is informed that knowledge of the mode of action of induction of tumors in the relevant dose range could inform conclusions on human relevance. Conduct of appropriate studies to address important data needs and uncertainties in the mode of action analysis can then be considered collectively by the risk manager/risk assessor in a refined problem formulation, depending on resources available and time frame for completion.

If additional generation of data is deemed appropriate, the assessment enters the "research" portion of the roadmap, but with a focused effort on generating data relevant to the mode of action/risk assessment question at hand. The targeted relevant mechanistic data that would inform additional assessment and/or management do not require full knowledge of the mechanism, but rather often quantitative information on determinants of key events, as a basis to predict interspecies differences and human variability better. Upon completion of relevant studies and subsequent mode of action/species concordance analysis, the risk manager is informed of the conclusion (i.e., whether data are considered sufficient to support the hypothesis that the tumors are unlikely to be of relevance to humans).

A potential variant includes the scenario that since the initial problem formulation, the risk manager has become aware that several other related chemicals co-occur with the substance of interest, which may be appropriate for consideration in the same category with chemical X in the risk assessment. The risk manager is informed that the rationale for inclusion of other category members would be strengthened if the same mode of action was suspected; relative potency could then be considered through targeted testing of an early key event. The assessment process now enters the "assessment-specific data generation" portion of the roadmap. Problem formulation can be an iterative process; thus, the results of the targeted testing would further inform the risk manager as to which chemicals within the category are hypothesized to act via the same mode of action, and therefore which should be included for read-across in a combined risk assessment. The assessment process then enters the final "risk assessment" portion of the roadmap.

Modified Framework

The mode of action framework addresses two key questions. The first is whether there are sufficient data to hypothesize, with an acceptable level of confidence, a mode of action for a known or suspected toxicological outcome. The second is the extent to which such a mode of action would, or is likely to, operate in humans at relevant exposure levels (species concordance analysis).

The framework can also be used in two quite different ways, the first reflecting how it was initially developed, for relatively data-rich chemicals. In this case, causal key events related to an observed (adverse) effect associated with a specific chemical exposure are

identified as a basis to utilize available data on kinetics and dynamics maximally to inform relevance to humans and subsequent dose–response analysis; this is referenced below as "Application of the mode of action framework for observed (adverse) effects" and reflects historical experience as is illustrated in many of the case studies currently available. Following problem formulation (Figs 2 and 3), then, a decision may be taken that a mode of action analysis would be of value in addressing an observed toxicological response for which the margin between measures of hazard and estimated human exposure is such that it warrants additional refinement of the assessment.

The second way in which the framework can be applied is based on information on key events from appropriate *in vitro* and *in silico* systems to predict and assess potential modes of action and potential consequent (adverse) effects (referenced below as "Application of the mode of action framework in hypothesizing (adverse) effects"). The outcome of such an analysis may be the development of a plausible case to predict an (adverse) effect based on knowledge of putative key events or, alternatively, the probable exclusion of certain (adverse) effects, based on an absence of a likelihood of perturbation leading to relevant key events.

In this context, mode of action comprises a series of causally associated key events leading to, potentially leading to or hypothesized to lead to an (adverse) effect. Hence, there can be only one mode of action for one chemical or group of chemicals leading to a specified effect under a given set of conditions. However, different chemicals, or the same chemical under different conditions (e.g., at higher doses or concentrations), may produce the same effect via different modes of action. An example would be the generation of site of contact tumors in the nasal cavity. One chemical may produce such an effect through cytotoxicity and subsequent cell replication promoting spontaneous mutations, another through DNA reactivity leading to gene mutations promoted by regenerative proliferation secondary to cytotoxicity, and a third through interaction with DNA leading to early mutations. In addition, early key events in competing pathways may, or often, converge to produce the same late key event (and outcome). Each mode of action comprising a series of key events for a given response will be different, but some of the key events may be common to other modes of action leading to the same response. The nature of the key events involved will have an impact on the shape of the dose–response curve and on interspecies and intraspecies differences.

The modified mode of action framework is outlined in Fig. 4 and explained in further detail below.

Application of the Mode of Action Framework for Observed (Adverse) Effects

Only this first approach was addressed in the previous descriptions of the WHO/IPCS/ILSI-RSI mode of action/human relevance framework (Boobis *et al.*, 2006, 2008; Meek *et al.*, 2003; Seed *et al.*, 2005), from which further detailed information can be obtained. Extension of the approach through application to help construct more predictive groupings of chemicals was subsequently highlighted in Carmichael *et al.* (2011). A key aspect of the approach, as illustrated through case studies, is that there should be an unequivocal effect to address before embarking on a mode of action analysis. Hence, problem formulation will

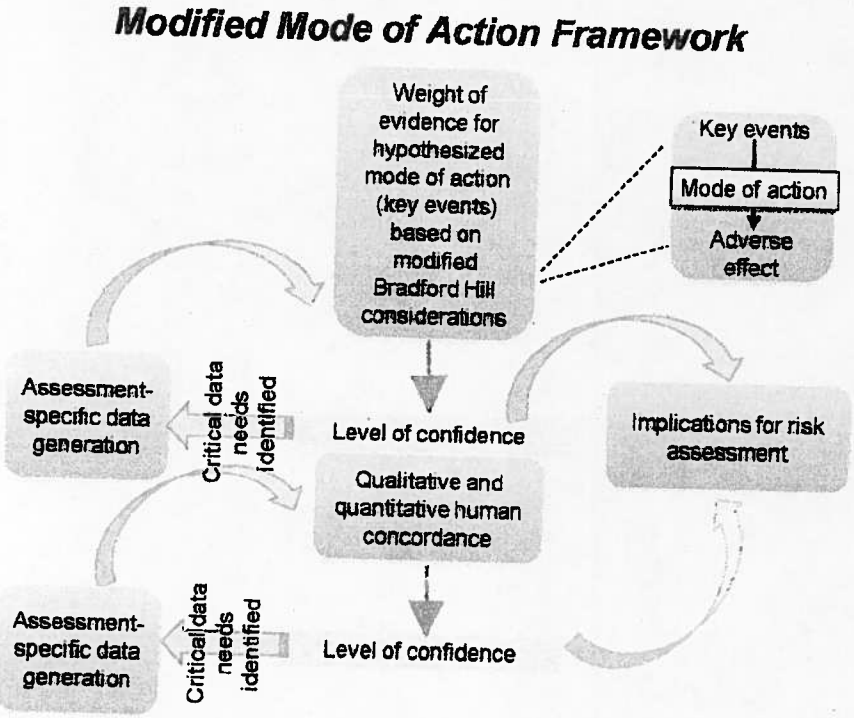


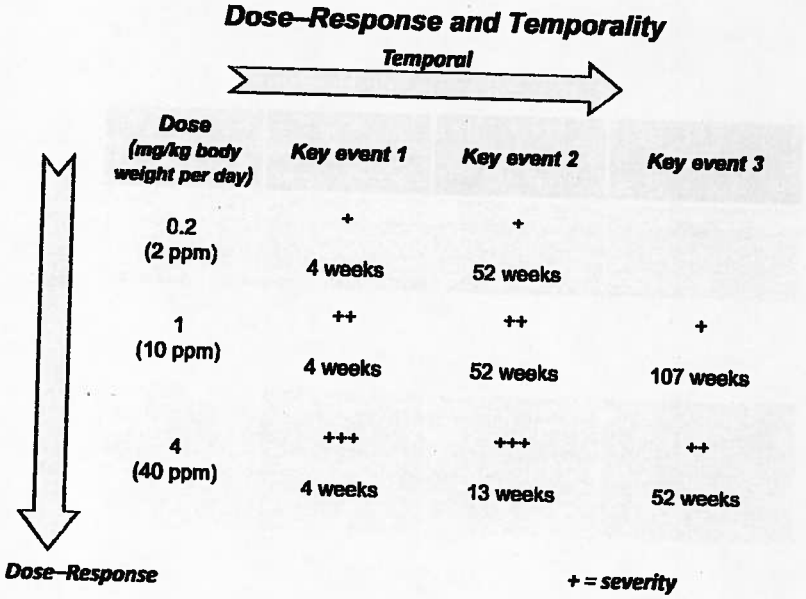
Figure 4. Modified mode of action/human relevance framework and its relation to data needs identified and risk assessment. The application of the framework to assess for observed (adverse) effects and in hypothesizing (adverse) effects is illustrated. The iterative nature of the analysis and the importance of expressing uncertainty are also highlighted.

have identified the (critical) effect(s) of concern to be considered in the analysis.

In general, mode of action analysis applies to a single effect in a single tissue. In essence, there is one mode of action leading to an effect of interest in the relevant organ for a given substance. This mode of action entails several key events, each of which may result from different (sometimes) competing mechanisms and/or pathways, although these converge at a late stage to produce the (adverse) effect. It is important, then, to robustly synthesize available information based on multidisciplinary input in hypothesizing potential modes of action. In addition, in the absence of information to the contrary, site concordance between animals and humans is generally assumed, at least as an initial premise. This is often the case, for example, for many non-genotoxic carcinogens that act through perturbation of physiological processes. Similarly, for many non-cancer

Modified Bradford Hill Considerations

- Concordance of dose–response relationships between key and end events
  - Dose–response relationships for key events would be compared with one another and with those for endpoints of concern
    - Are the key events always observed at doses below or similar to those associated with the toxic outcome?
- Temporal association (time)
  - Key events and adverse outcomes would be evaluated to determine if they occur in expected order



- Consistency and specificity
  - Is the incidence of the toxic effect consistent with that for the key events?
    - i.e., Less than that for the key events?
  - Is the sequence of events reversible if dosing is stopped or a key event prevented?
- Biological plausibility
  - Is the pattern of effects across species/strains/systems consistent with the hypothesized mode of action?
  - Does the hypothesized mode of action make sense based on broader knowledge (e.g., biology, established mode of action)?

Figure 5. An illustration of the modified Bradford Hill considerations for weight of evidence of hypothesized modes of action. The illustration represents evolution of these considerations based on increasing experience in application in case studies and training initiatives internationally. Specific questions being addressed by each of the considerations are offered as a basis potentially to increase common understanding and consistency in their application in mode of action analysis.

Comparative Weight of Evidence for Hypothesized Modes of Action: Cytotoxic Mode of Action Example Summaries

Modified Bradford Hill consideration	Supporting evidence	Potentially inconsistent evidence
Dose–response Temporal concordance	Metabolism, cytotoxicity, proliferation precede tumors; tumors observed only at cytotoxic doses (benchmark dose analysis) *(quality based on nature and number of studies)*	Tumors observed at doses lower than those at which key events observed
Consistency, specificity	Consistency in repeated studies and different labs and across species, sexes, routes and levels of biological organization *(#s)* correlating with extent of metabolism. No adverse effects without relevant enzyme in null mice. Incidence of tumors less than that for key events and tissue recovery in reversibility studies.	Incidence of tumors greater than that for key events
Biological plausibility	Consistency with state of knowledge on cancer	

Comparative Weight of Evidence for Hypothesized Modes of Action: Mutagenic Mode of Action[a] Example Summaries

Modified Bradford Hill consideration	Supporting evidence	Potentially inconsistent evidence
Dose–response Temporal concordance	Dose–response and temporal pattern for genotoxicity and tumors consistent with the compound acting via a mutagenic mode of action	Parent compound negative for mutation in a range of *in vitro* and *in vivo* bioassays *(quality based on nature and number of studies)*
Consistency, specificity	Evidence in a range of well conducted bioassays that mutation is an important early key event (e.g., occurs early and at relevant doses)	The pattern of genotoxicity results inconsistent with what would be expected for the hypothesized mode of action (e.g., not mutagenic in a range of assays; metabolite induces mutation at cytotoxic doses)
Biological plausibility	Pattern of results for genotoxicity consistent with that observed for chemicals known to act via a mutagenic mode of action	Pattern of results for genotoxicity inconsistent with that observed for chemicals known to act via a mutagenic mode of action

[a]Where mutation is an early and influential key event.

Figure 6. An example of comparative weight of evidence for hypothesized cytotoxic and mutagenic modes of action. Information in each of the columns provides an overview of the extent and nature of the available data and its cohesiveness. Particularly important in interpretation of relative weight of evidence is the nature and extent of data that may be inconsistent with a hypothesized mode of action. In this particular case, the extent of inconsistent data is considerably less for a hypothesized mode of action where mutation is likely to be secondary to cytotoxicity than for a mutagenic mode of action (i.e., where mutation is an early and influential key event). Indeed, the pattern of data on genotoxicity is completely consistent with a cytotoxic mode of action. This would lead to the conclusion that there is greater confidence in the chemical acting by a cytotoxic than by a mutagenic mode of action.

Concordance Table with Dose–Response

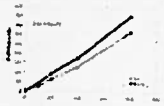
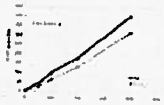
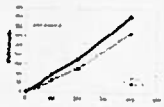
Key event / adverse outcome	Qualitative species concordance	Evidence base	Quantitative species concordance	Quantitative dose–response
Metabolism by cytochrome P450 2E1	Relevant enzyme in kidney and liver of humans	Considerable in animals; limited but relevant to humans	Physiologically based pharmacokinetic model incorporating metabolic rates, enzyme affinities and distribution based on *in vitro* human data supported by *in vivo* data	
Sustained cell damage and repair (cytotoxicity, proliferation)	Liver and kidney target organs in humans	Considerable in animals; possible in humans, but limited data	No data	
Liver and kidney tumors	Possible in humans	Considerable in animals; highly plausible in humans	No data	

Figure 7. An illustration of a concordance table including dose–response curve. The kinetic and dynamic data considered in assessment of mode of action are directly relevant to dose–response analysis, which takes into consideration dose–response relationships for each of the key events.

endpoints, site concordance between test species and humans is a reasonable first assumption, based on considerations of biological plausibility and chemical-specific mechanistic data.

However, there are exceptions to this general principle. Consistent with species- and tissue-specific variation in metabolic activation and detoxification, site concordance for DNA-reactive carcinogens or other effects for which metabolism is critical is often poor. Similarly, for some non-cancer effects induced through a pleiotropic response, such as those that are endocrine mediated, site concordance should not be assumed, but rather considered, based on available mechanistic data and knowledge related to biological plausibility.

These possibilities would need to be scoped at the outset of any mode of action analysis. In such cases, it may be that mode of action analysis would benefit from considering multiple sites in the same evaluation. However, care must be taken to ensure that the mode of action for each effect is likely to be the same, which will not always be the case.

Mode of action analysis relies upon biological plausibility and coherence. The weight of evidence for a hypothesized mode of action is addressed based on the Bradford Hill considerations, proposed originally to examine causality of associations observed in epidemiological studies, but later modified in WHO/IPCS and ILSI-RSI publications on the mode of action/human relevance framework (Boobis *et al.*, 2006, 2008; Meek *et al.*, 2003; Seed *et al.*, 2005) and additionally evolved, here. The original templates for consideration of the weight of evidence for a hypothesized mode of action were based on consideration of traditional measures of toxicity, such as biochemical and histopathological parameters in experimental animals. These templates have been adapted here (Figs 5–7) to reflect additional experience gained in the application of the framework in an appreciable number of case studies over the past decade and as a basis potentially to encompass additional early key events from evolving methods to reliably predict human health outcomes. Based on this experience, robust consideration of dose–response relationships and temporal concordance for early key events will be important in documenting weight of evidence for proposed adverse outcome pathways.

Relevant considerations include dose–response relationships and temporal concordance between specified key events and outcome, consistency (of, for example, the incidence of key events and outcome and changes in causally associated key events), specificity (in the context of essentiality of key events and reversibility) and biological plausibility, based on coherence with the state of knowledge.

In relation to dose–response relationships and temporal concordance, a key event cannot play a role in an (adverse) effect if it is manifest only after toxicity has occurred or if it occurs only at doses higher than those inducing toxicity. The same applies to late key events relative to early key events. There is often a close relationship between dose and time dependency, so that the higher the dose, the earlier a key event is observably affected, and vice versa. This pattern of dose–response and time–response relationships can be invaluable in assessing weight of evidence for a hypothesized mode of action and its key events or how different key events are interrelated. Systematic consideration of dose–response relationships and temporal concordance between key events and (adverse) effects, as illustrated in Fig. 5, encourages early assimilation of relevant information from the broader database of both short- and long-term studies, or from different non-animal test systems, in a mode of action context.

More detailed discussion on all of the modified Bradford Hill considerations when applied in the mode of action analysis for observed (adverse) effects is provided in previous publications on the mode of action/human relevance framework and will not be repeated here. Application and weighting of these considerations continue to evolve as a basis to additionally increase consistency and transparency in assessing weight of evidence in mode of action/species concordance analysis.

It is essential at the outset of mode of action/species concordance analysis that all reasonably plausible modes of action be considered. These include those modes of action that have previously been associated with the relevant effect and any series of key events that logically presents because of available experimental information. The case for each plausible mode of action should be evaluated systematically from the outset, using modified Bradford Hill considerations.

Weight of evidence for alternative hypotheses should be considered and assessed comparatively. Figure 6 illustrates such an evaluation. Based on relative weight of evidence, it can be determined whether one mode of action could be considered with reasonable certainty to explain the (adverse) effect. Where it is not possible to exclude one or more modes of action, critical data needs could be identified as a basis to inform relevant research that could reduce uncertainty concerning the causal key events within a mode of action, depending on the needs and urgency of the assessment as considered in problem formulation.

The degree of confidence in the outcome should be specified, and each step in the mode of action analysis should be accompanied by a list of the critical uncertainties (i.e., lack of knowledge) and associated data needs, prioritized on the basis of their likely impact, if filled, on weight of evidence and implications for subsequent dose–response analysis.

The comparative analysis of weight of evidence for hypothesized modes of action based on the modified Bradford Hill considerations is followed by statements on the likelihood of each being operative to induce the critical effect. Alternatively, depending on the needs and urgency of the assessment addressed in problem formulation, plausible modes of action should be considered as a basis to contrast strengths and weaknesses of different approaches to quantification of interspecies and intraspecies extrapolation in dose–response modeling. This enables risk managers to distinguish best-supported options (i.e., those that are most certain), which is critical in increasing transparency in separating science judgment (i.e., considerations based on experienced consideration of the relevant science base) from science policy determinations (e.g., embedded conservatism in human health risk assessment, incorporated to increase public health protection). Characterization of this nature also contributes to consistency across weight of evidence considerations in different mode of action analyses.

An important objective of framework analysis, then, is the description of the critical sources of uncertainty and characterization of their impact on conclusions concerning weight of evidence for various hypothesized modes of action and their relevance to humans, as a basis particularly for identification of priorities for generation of more or better data. Sensitivity of the estimate to various assumptions can also be tested, and/or available quantitative data relevant to key uncertainties can be analyzed.

Following mode of action analysis and consideration of the associated uncertainties, several outcomes are possible, as illustrated in Fig. 4. There may be sufficient information to conclude that a hypothesized mode of action is supported by available evidence to explain the effect of concern and that key events for this mode of action have been clearly identified. Where there is insufficient information to reach a conclusion with adequate confidence that a hypothesized mode of action explains the (adverse) effect of concern, appropriate research to address identified critical data needs should provide suitable information to enable confirmation or otherwise of the hypothesized mode of action, through iterative application of the framework. Finally, it may be that at the conclusion of the analysis a hypothesized mode of action is rejected and no other mode of action logically presents itself. In such instances, it may

be necessary to proceed with the risk assessment empirically, using relevant information that has been obtained during the analysis of the mode of action – for example, dose–response and time–response information on the endpoint itself, or relevant kinetic and dynamic data.

An important objective of mode of action analysis is to identify those key events that are likely to be most influential in determining potential qualitative and quantitative differences within and between species – that is, key events that are dose and rate limiting. This is addressed in species concordance analysis and is illustrated in Fig. 7. Where it has been possible to conclude that a hypothesized mode of action is adequately supported by the available information with an acceptable level of confidence, it is necessary to consider the extent to which such a mode of action would, or is likely to, operate in humans. Species concordance analysis starts with a statement on the level of confidence in the weight of evidence for the hypothesized mode of action under consideration and associated uncertainties. The extent of this analysis is necessarily dependent upon the test system(s) in which key events have been measured, being less for those that best represent humans.

Consideration of mode of action also enables identification of early events or indicators of susceptibility that could be measured in humans (i.e., biomarkers); for example, if there is sufficient information to support early key events such as metabolic activation to a reactive metabolite, this directs attention to the relevant parameters in humans, as a basis to predict interspecies (based on comparison of the relevant parameters between humans and animals, scaled as appropriate) and intraspecies differences (based on consideration of the relevant parameters within different subgroups of the population). Consideration of potential key events also contributes to identification of any specific subpopulations (e.g., those with genetic predisposition or life stage differences) that may be at increased risk.

Assessment of concordance is accomplished by systematic consideration of the nature of the key events between and within species, taking into account both chemical-specific and more generic information, such as anatomical, physiological and biochemical variations. Concordance is considered both qualitatively and quantitatively (Fig. 7). On rare occasions, it may be possible to conclude that a mode of action identified in studies in animals is not relevant to humans because of profound qualitative differences identified in experimental investigation; for example, the molecular target necessary for a key event is not present in humans, and there is no functional equivalent. An example would be α_{2u}-globulin, which plays a key role in the renal carcinogenicity of D-limonene (see Case example 1) (Meek *et al.*, 2003). Alternatively, and very infrequently, quantitative differences in key events may be so great as to render the mode of action not relevant to humans at any conceivable exposure to the substance.

Case example 1: Lack of human concordance

D-Limonene provides an example of a data-rich case example for which the mode of action has been established with confidence in the animal model and extensive data are available to demonstrate that it is not relevant to humans (Meek *et al.*, 2003).

Hypothesized key events in the mode of action for species- and sex-specific kidney tumors in male rats were the formation of a stable intermediate, D-limonene-1,2-epoxide, which binds to a protein, α_{2u}-globulin, which accumulates in the renal proximal tubule cells, leading to nephropathy and cellular proliferation, and subsequently tumors, at this site following chronic exposure. There is strong evidence that female rats, laboratory mice and other strains of rats for which there is no evidence of D-limonene-related renal toxicity or tumors do not synthesize or express α_{2u}-globulin.

Consideration of the relevance to humans of the key events leading to renal tumors in the male rat model identified the expression of either α_{2u}-globulin or a homologous protein in humans as critical. After an exhaustive analysis, no protein capable of binding to D-limonene-1,2-epoxide could be identified from human kidney, and therefore it could be concluded that the mode of action leading to kidney tumors in the male rat was not likely to be operable in humans.

This is a rare example of a distinct qualitative difference between the animal model and humans, allowing the possibility to conclude that a mode of action is not relevant to humans. However, it is quite unusual to be able to demonstrate such a qualitative difference. Rather, in the vast majority of cases, such differences will be quantitative, and likely differences in sensitivity of response between animals and humans identified in the mode of action analysis would be taken into account in the subsequent dose–response analysis.

If the weight of evidence for the hypothesized mode of action is sufficient and its relevance for risk assessment cannot be excluded, the implications for dose–response analysis and population variability are considered in the context of identified kinetic and dynamic data. Figure 7 indicates the relevance of delineation of key events in hypothesized modes of action considered to operate in humans in subsequent dose–response analysis. In fact, there is a dose–response curve for each of the key events, and risk for the human population is best predicted on the basis of those key events (or a combination thereof) that are likely to be most influential in impacting or preventing risk, taking into account potential interspecies and interindividual differences in kinetics and dynamics as considered in the species concordance analysis. Reliance on earlier key events offers the potential to better characterize and/or acquire data on effects at lower doses or concentrations in human tissues or populations, which are more relevant for risk assessment. It also contributes to the development of more relevant and informative data for human life stages and subpopulations. For Case example 2, these data could be used additionally in quantitative species concordance analysis, with implications for subsequent dose–response analysis, the identification of critical data needs and the contribution of evolving methods – in this case, well-designed genomic studies – see "Application of the mode of action framework in hypothesizing (adverse) effects" below (see also Table 2).

Case example 2: Use of kinetic and dynamic data in species concordance analysis and implications for dose–response analysis – Contribution of well-designed genomic studies

This example illustrates the manner in which kinetic and dynamic data may potentially inform quantitative concordance analysis, including interspecies variation and human

variability and, subsequently, dose–response analysis and extrapolation. The example also illustrates how mode of action/species concordance analysis informs meaningful generation of critical data relevant to risk assessment, including that from evolving methods.

Cacodylic acid (dimethylarsinic acid) is a pesticide that causes dose-related increases in the incidence of bladder tumors in rats, but not mice (Cohen *et al.*, 2006b, 2007; US EPA, 2005b). Incidence is increased significantly only at the highest administered dose levels. The parent compound undergoes reductive metabolism to a toxic metabolite, and observed damage to urinary epithelial cells correlates with this pathway (see Cohen *et al.*, 2006b; US EPA, 2005b). The levels of toxic metabolite are significantly increased at doses causing cytotoxicity, proliferative regeneration and bladder tumors. The weight of evidence from critically evaluated data from a wide range of assays both *in vitro* and *in vivo* indicates that the parent compound is not mutagenic, but that the active metabolite is clastogenic at high concentrations or doses. The concentration–response relationships for cytotoxicity associated with the active metabolite were similar in *in vitro* studies in bladder cells of rats and humans. Because of toxicokinetic differences, the toxic metabolite is expected to form at a lesser amount in human urine compared with rats (Cohen *et al.*, 2006b; US EPA, 2005b).

Application of the modified Bradford Hill considerations supported the weight of evidence for the hypothesized key events in the mode of action, which included reductive metabolism and cytotoxicity and proliferative regeneration leading to bladder tumors (Cohen *et al.*, 2006b; US EPA, 2005b). Weight of evidence considerations included a thorough analysis of dose–response relationships and temporal concordance as determined from benchmark dose analyses of a range of *in vivo* studies of different durations. This does not imply a 1:1 correlation of the incidence of early and late key events (rather, the incidence of early key events is expected to be higher), as key events are essential, but not necessarily sufficient in their own right.

Qualitative and quantitative concordance analysis based on relevant kinetic and dynamic data indicated that these effects are relevant to humans and that quantitative differences would most likely be related to extent of delivery to the target organ of the toxic metabolite and variations in sensitivity of the bladder to damage induced by this metabolite. Chemical-specific adjustment factors could then be derived from a physiologically based pharmacokinetic model incorporating metabolic rates, enzyme affinities and distribution based on *in vitro* human data supported by *in vivo* data and quantitative reflection of the similarity in sensitivity to the active metabolite between the rat and human bladder in *in vitro* studies.

The mode of induction of bladder tumors was deduced principally based on key cytological and biochemical events in mechanistic studies from experiments designed to address critical aspects of both the mode of action and species concordance analysis. The results of genomic studies indicated that similar networks were altered in rat and human urothelial cells exposed to the active metabolite at doses similar to those in urine at which tumors were observed in the critical bioassays. The concordance table in Table 2 outlines confidence/uncertainties in the mode of action/species concordance analysis.

Mode of action analysis also contributes to the interpretation of relatively extensive epidemiological data sets. For example, information on key events in mechanistic studies can contribute to better understanding of expected (not necessarily similar) target organs in humans. This is relevant to the interpretation of negative epidemiological data based on their power to detect the most likely site of damage in humans taking into account mode of action and interspecies differences in key determinants of key events. It also contributes to the selection of appropriate biomarkers of effect in epidemiological studies and to understanding of variations between life stages and subgroups of the human population (see Case example 3).

Case example 3: Role of mode of action analysis in the evaluation of epidemiological data

This case example illustrates the contribution of mode of action analysis when there is substantial human evidence.

Associations between ambient particulate matter exposures and increased cardiovascular mortality were first observed in epidemiological studies without support from animal bioassays, which led to skepticism concerning causality due to the lack of mechanistic underpinning. Subsequent mode of action studies shed light on key events in cardiovascular injury in humans exposed to particulate matter and elucidated interspecies differences and human variability in dosimetry and sensitivity (US EPA, 2009b).

Particulate matter induces adverse effects on the cardiovascular and cerebrovascular systems, such as thrombosis, plaque rupture, myocardial infarction and stroke, via reactive oxygen species, which appear to trigger systemic inflammation through the action of cytokines and other soluble mediators. In general, systemic inflammation is associated with changes in circulating white blood cells, the acute phase response, procoagulation effects, endothelial dysfunction and the development of atherosclerosis. The time course of these responses varies according to the acute or chronic nature of the particulate matter exposure; chronic exposures may also lead to adaptive responses.

If there is appreciable uncertainty about the relevance or applicability of a mode of action, but critical data needs can be identified, it may be possible to obtain such information through conduct of appropriate studies. Table 2 includes the concordance analysis for the example included in Case example 2, illustrating principal areas of uncertainty, where generation of additional data might meaningfully inform the risk assessment.

If it is not possible to establish whether a mode of action would, or is likely to, operate in humans with an acceptable level of confidence, but there is a pressing need for risk management decisions because of the urgency or the nature of the problem, knowledge of dose–response relationships and variability across species may still be of value in later stages of the risk assessment.

The conclusions of the concordance analysis should be accompanied by consideration of associated uncertainty and a statement on the level of confidence that a mode of action would, or is likely to, operate in humans.

J. Appl. Toxicol. 2014; **34**: 1–18

wileyonlinelibrary.com/journal/jat

Table 2. Concordance analysis of key events in the mode of action associated with induction of bladder tumors in rats by cacodylic acid (Cohen et al., 2006b; US EPA, 2005b)

Key event	Qualitative concordance		Quantitative concordance	Confidence/uncertainty
	Rats	Humans		
Reduction of cacodylic acid (dimethylarsinic acid, or DMA^V) to the highly cytotoxic metabolite, dimethylarsinous acid (DMA^{III}), in urine	Yes: *in vivo* studies detecting DMA^{III} in urine at concentrations that would produce cytotoxicity after DMA^V is administered.	Plausible: Evidence following DMA^V exposure too limited to draw conclusions, but DMA^{III} shown to be present following human exposure to inorganic arsenic.	Formation of less DMA^{III} in urine of humans compared with rats. Significant levels of additional metabolite trimethylarsine oxide (TMAO) in rodents; detected in humans only at very high doses of inorganic arsenic. DMA^V is a poor substrate for the arsenic(III) methyltransferase (AS3MT) in humans. Variation between humans and rats in transport of DMA^V across cell membranes. Similar magnitude of response of human and rat epithelial cells to DMA^{III}. Interspecies differences could be taken into account in dose–response analysis through physiologically based pharmacokinetic modeling and use of chemical-specific adjustment factor for dynamics.	Considerable evidence in animals; limited in humans.
Urothelial cytotoxicity	Yes: Scanning electron micrographs of rat urothelium; *in vivo* cytotoxicity findings correlate closely with *in vitro* studies.	Human evidence from *in vitro* studies of urothelial cells, potential to occur *in vivo* in humans if sufficient DMA^{III} is formed.		Considerable consistent evidence that the metabolite leading to urothelial cytotoxicity is DMA^{III} and that cytotoxicity is a rate-limiting key event; quantitative species differences in key events (mode of action) can be taken into account.[a]
Regenerative urothelial proliferation	Yes: *In vivo* 5-bromo-2′-deoxyuridine labeling index data.	No human evidence, but potential to occur in humans if sufficient cell killing is produced and sustained.		Considerable evidence in animals, although some inconsistencies in the data that can be accounted for by variability across different laboratory studies.
Development of urothelial tumors	Yes: Responses in rats but not mice.	No epidemiological data: Only if humans were exposed to doses of DMA^V that are sufficiently high to lead to cytotoxic levels of DMA^{III} in the urine.		Strong and consistent evidence supporting the sequence of key events postulated for the development of rat bladder tumors. Good understanding of species differences impacting key events. Evidence in humans is weak. Mode of action is qualitatively plausible in humans, presuming sufficient DMA^{III} is present in the urine.

[a]Though the biochemical target for cytotoxicity is not understood, this information is not essential for the mode of action.

Application of the Mode of Action Framework in Hypothesizing (Adverse) Effects

Lessons learned in mode of action/species concordance analysis for identified effects are also relevant to its application where the (adverse) effect is not demonstrated but could potentially be presumed based on measurement of putative early key events in established modes of action, taking into account lines of available evidence.

Thus, hypotheses about the key events that can lead to the observed (adverse) effect of concern are developed. In contrast, one can also develop hypotheses of potential (adverse) effects that may be triggered by observed putative early key events, based on previous generic knowledge on documented modes of action. Both approaches involve an iterative process of hypothesis testing and data generation.

In this approach, the objective is to identify those modes of action that could plausibly arise from the (series of) key events identified, either because of previous knowledge of their involvement in a mode of action (e.g., for related chemicals for which there are more data) or because a plausible case can be made on the basis of existing biological understanding that such (a series of) events or perturbations may reasonably lead to (adverse) outcomes under certain time- and dose-dependent conditions. The methods used for evaluating putative modes of action will be fit for purpose, which will not necessarily involve one-for-one validation against existing *in vivo* methods. Thus, at the outset, consideration of potential key events in the mode of action plays an integral role both in the choice of experimental methods (*in vivo*, *in vitro* or *ex vivo*) and in data interpretation. Based on the understanding of the causal linkage of putative key events (either observed or anticipated), hypotheses of the likely potential effects of exposure to a chemical are developed in mode of action analysis. Thus, the modified Bradford Hill considerations are just as applicable here, but are not yet well tested.

In terms of quantitative dose–response assessment of the key events, a critical factor is extrapolation of the effect levels *in vitro* or predicted *in silico* to target tissue concentration *in vivo* – for example, by using physiologically based toxicokinetic modeling (referenced as quantitative *in vitro* to *in vivo* extrapolation modeling). Thus, a key consideration is target tissue concentration of the toxicologically active moiety. This approach lends itself well to identification of the causative agent (i.e., parent or metabolite) and readily enables qualitative and quantitative information to be obtained on the enzyme reactions involved. It may be possible to discount human relevance of some putative modes of action based on the margin between effect levels *in vitro* and anticipated target tissue concentrations *in vivo*. This may be particularly important in the short term, when there is substantial uncertainty about the significance of weak signals obtained using *in vitro* methods.

As discussed above, confidence in a mode of action postulated on the basis of putative early key events identified using non-animal methods will depend on the weight of evidence linking these key events with a mode of action for an adverse response from previous studies and on the ability to "calibrate" quantitative changes in the key event against a degree of change known to have adverse consequences. An example would be inhibition of an enzyme involved in neurotransmitter synthesis or degradation. The extent to which this enzyme needs to be inhibited to produce adverse consequences may be known from studies *in vivo* and could then be used to calibrate such changes determined *in vitro* or predicted *in silico*. Integral to this would be knowledge of the extent to which adaptive mechanisms operating *in vivo* are functional *in vitro* or included in the *in silico* model systems.

Formal analysis of site concordance for key events may not be necessary in this approach. Similar to the mode of action analysis for observed (adverse) effects, data may have been generated in tissue-specific model systems or may reflect site-specific key events. Prediction of likely site of effect will require additional considerations, such as the uptake and disposition of the chemical and the activity of causal pathways in different tissues and cell types. For example, if toxicity depends in part upon transport into the target cell to reach a critical concentration, the presence of the transporter in different cell types would be a key consideration in assessing potential site specificity. Similarly, if one of the key events involved inhibition of a specific potassium channel, the tissue distribution of this ion channel would be an important factor in assessing site specificity. Eventually, as knowledge of the biology of the causal pathways increases, it may be possible to use a systems approach to predict likely affected tissues.

Critical to interpretation of data obtained using non-animal methods will be the model system in which information on putative early key events was obtained and whether coverage of more than one key event would be expected. Some key events may be assessed individually (e.g., using *in silico* approaches to predict binding affinity to a receptor), whereas others may be assessed in a more integrated system (e.g., cytotoxicity in a metabolically competent cell system). Alternatively, high-content analysis and bioinformatics may be used to identify those pathways affected by a substance.

In the case of a well-established mode of action, the focus is on determining whether the measured key events provide sufficient evidence to accept the plausibility for the (adverse) outcome without necessarily generating *in vivo* data specifically to demonstrate the (adverse) outcome. Where the mode of action has not previously been established, the possibility that a plausible case can be made because of existing biological understanding should be addressed. Failing this, the likely outcome of such an analysis is the generation of a hypothesis for a possible (adverse) effect, which can then be tested *in vivo*. In any event, once a mode of action is established, the key events are known a priori and can then be assessed *in vitro* or *in silico*, Thus, by understanding the likelihood of effects (i.e., initiation of a toxicity pathway) at lower levels of biological organization (e.g., from SARs and *in vitro* models), it can be determined if more expensive and time-consuming testing at higher levels of biological organization (i.e., *in vivo*) is needed, contributing to increasing efficiency in hazard testing of chemicals. Viewed from the opposite perspective, certain *in vivo* testing could be eliminated for substances that show no potential to initiate the chain of events comprising the mode of action for an (adverse) outcome at environmentally relevant concentrations. In other words, tailored testing can be developed according to screening outcomes indicating the potential for (adverse) effects (see Case example 4).

J. Appl. Toxicol. 2014; 34: 1–18

Case example 4: Use of mode of action analysis to guide development of more efficient testing strategies

Concepts of mode of action analysis are also helpful in guiding developments in the replacement of *in vivo* toxicity testing.

Modes of action can be hypothesized based on reference chemicals/pharmaceuticals where the sequence of key events leading to a specific (adverse) effect is known at a sufficient level of detail, as a basis to facilitate identification of the characteristics and requirements of *in vitro* systems and *in silico* models that could predict early and subsequent rate-limiting key events in an integrated manner. Once dose–response relationships between the key events measured *in vitro* and biomarkers of response and ultimately adverse outcome *in vivo* are established for reference chemicals, including the necessary *in vitro* to *in vivo* extrapolation, the toxicity of many other chemicals acting through the same mode of action could in theory be characterized and predicted based on the responses in the *in vitro* systems and *in silico* models.

A large research initiative ("Safety Evaluation Ultimately Replacing Animal Testing," or SEURAT) is based on this premise (Gocht *et al.*, 2013). The first phase of this program, which is co-funded by the European Commission under its Seventh Framework Programme (FP7) and Cosmetics Europe, spans a 5-year period from 2011 to 2015 and includes six research projects, combining the research efforts of over 70 European universities, public research institutes and companies addressing repeated-dose toxicity in hepatic, cardiac, renal, neuronal, muscle and skin tissues. The strategy involves mode of action analysis to describe how any substance may adversely affect human health and to use this knowledge to develop complementary theoretical, computational and experimental (*in vitro*) models that predict quantitative points of departure for safety and risk assessment.

Where data are available on only one or a limited number of key events and the link to an (adverse) effect has not been sufficiently demonstrated, the data may still be of value in helping to rank and prioritize chemicals, as a basis for additional testing and/or decision-making based on likely relative hazard (e.g., relative potency in modulating sodium channels, endocrine disrupting substance prioritization) (see Case example 5).

Case example 5: Mode of action analysis in prioritizing substances for further testing

There is a great deal of interest in prioritizing chemicals for evaluation of endocrine disruption potential (i.e., how best to focus on those chemicals most likely to cause adverse effects without empirically testing all chemicals of regulatory concern). An expert (QSAR) system was developed to predict estrogen receptor binding affinity, using the mode of action (adverse outcome pathway) knowledge (OECD, 2009; Schmieder *et al.*, 2003, 2004; US EPA, 2009a). This pathway is initiated through direct chemical binding to the estrogen receptor, which could plausibly lead to reproductive impairment. The predictive model was developed based on two *in vitro* assays: using a rainbow trout estrogen receptor competitive binding assay to directly measure the chemical–biological interaction and a trout liver slice assay in which the consequences of estrogen receptor activation or inhibition are measurable as a result of tissue uptake and partitioning of the chemical in the presence of xenobiotic metabolism.

More broadly, consideration of SARs for specific key events known to be involved in the mode of action of representative chemicals with the same structural features would be invaluable in helping to construct chemical categories and would enhance the reliability of read-across (see Case example 6 on pyrethroids and Case example 7 on aniline).

Case example 6: Mode of action in the creation of chemical categories

This example addresses the risk assessment of a new synthetic pyrethroid with the same pesticidal mode of action and insecticidal effects as other members of this structural class of compounds. The critical effect of most pyrethroids is reversible neurotoxicity through interaction with a common target, neuronal sodium channels (reviewed in Soderlund, 2012). This mode of action has been established with confidence, and hence the similarity of the pesticidal mode of action of a new member of this chemical group will provide evidence that the compounds share key events. This can be used to support read-across. The risk assessment of a new pyrethroid could then be based on the assumption that it will share a mode of action with other pyrethroids and its likely relative hazard considered in this manner for a first-tier assessment.

The mode of action involves interaction with neuronal sodium channels (Clark and Symington, 2012; Soderlund, 2012). Hence, interaction with sodium channels is a key event for what is often the critical effect. One could rank existing pyrethroids for their potency in modifying the neuronal sodium channel in a suitably designed *in vitro* system and determine the potency of the new compound in this system (Cao *et al.*, 2011b; McConnell *et al.*, 2012). One would also wish to consider basic toxicokinetic aspects, such as absorption (which could be predicted from lipid solubility) (Hou *et al.*, 2009) and metabolic stability (which could be determined in *in vitro* test systems, such as hepatic microsomal fraction or cultured hepatocytes) (Scollon *et al.*, 2009). This information could be used, either semiquantitatively or with a physiologically based toxicokinetic model (Knaak *et al.*, 2012), to inform the choice of reference point from among those of the compounds for which information is already available.

Hence, by using an established mode of action for a structurally well-defined group of compounds with a common toxicophore, it is possible to inform read-across in the early tiers of a risk assessment. This could be refined by evaluating specific key events *in vitro* and using the resulting information to refine the read-across process. In this way, the results of new *in vitro* approaches can be anchored in relevant outcomes by using existing knowledge and concepts.

Journal of AppliedToxicology

In addition, such information would help in constructing assessment groups for consideration in the risk assessment of combined exposures to multiple chemicals (Cao *et al.*, 2011a).

Case example 7: Use of mode of action analysis to identify critical data needs and testing strategies in read-across

This case example is based on a case study presented at an Organisation for Economic Co-operation and Development (OECD) workshop held in December 2010. It addresses a mode of action related to the formation of methemoglobin and a number of industrial chemicals that are anilines, which vary in the quantity of toxicity data available (European Chemicals Bureau, 2004). It illustrates how the understanding of the mode of action can focus testing and more effectively fill data needs for data-limited compounds.

Aniline induces methemoglobinemia, which, if severe, can result in hemolytic anemia. Hemolytic anemia is a late consequence of methemoglobinemia and a response to the elimination of circulating red blood cells that contain methemoglobin. Aniline is first metabolized in the liver (probably by cytochrome P450 enzymes) to phenylhydroxylamine. It is further oxidized in red cells, most likely to free radical species, via nitrosobenzene. The iron in hemoglobin is oxidized by the free radical species from Fe^{2+} to Fe^{3+}, in which state (i.e., methemoglobin) it cannot bind oxygen. Decreased oxygen results in hypoxia-induced necrosis in tissues that have high oxygen needs. Damaged red blood cells are sequestered by the spleen and are phagocytosed by splenic macrophages, leading to increased red blood cell production by the blood-forming organs, primarily the bone marrow. If the bone marrow cannot keep up with the replacement needs, then extramedullary hematopoiesis occurs as a compensatory response. To determine the potential of the untested anilines to result in hemolytic anemia, *in vitro* testing could be conducted to measure the formation of phenylhydroxylamine and/or methemoglobin.

Thus, the mode of action framework provides a conceptual construct to consider key events at different levels of biological organization plausibly linked to an *in vivo* endpoint of regulatory interest. This allows for the development and use of alternative (*in vitro*) assays to target particular cellular or physiological key events along a specific pathway. Once the mode of action has been established, the key event data can be used for read-across from other chemicals. If a new chemical fits the established mode of action, this existing knowledge can be used to justify a more efficient testing strategy, so not every chemical needs to be evaluated in an *in vivo* test.

Information on mode of action, or on critical key events, can also be invaluable in helping to construct assessment groups for conducting a risk assessment of combined exposure to multiple chemicals (Meek *et al.*, 2011; see Case example 6).

One conclusion from the application of the mode of action framework to information obtained using non-animal methods could be that the data are sufficiently robust to support an established mode of action with a known causal relationship to an (adverse) outcome. Alternatively, it may be possible to conclude that whereas information on one or more key events is missing, provision of information on this data gap would enable a putative mode of action to be assessed with confidence. Finally, the available data may be such that it is not possible to postulate any mode of action with an acceptable degree of confidence.

Increasing numbers of data warehouses comprising substantial amounts of curated information on interspecies and interindividual variability in parameters relevant to many key events are becoming available. These warehouses cover a wide range of species- and individual-specific information, including human demographics, anatomical, physiological, biochemical, clinical chemical and life stage-dependent parameters, genetic, genomic, epigenetic, transcriptomic, proteomic and metabolomic information, phenotypic variation in cellular and physiological functions, and expression levels and activities of enzymes and transporters of xenobiotic disposition. Such information, together with evolving bioinformatics and computational tools, may facilitate quantitative (both deterministic and probabilistic) analyses of variability and more robust uncertainty analyses. These tools may also enable more effective analysis of the frequency with which alterations of key events and pathways are reported in similar studies, within and across animal species, and among humans. Similarly, they may permit a more thorough analysis of dose, exposure durations and response relationships in pathways across studies.

It should be noted that the availability of larger quantities of data on early potential key events to inform mode of action analyses might lend itself to probabilistic assessments and more robust uncertainty analyses.

Discussion and Conclusions

The WHO/IPCS mode of action/human relevance framework has been updated to reflect experience acquired in its application, as well as extending its utility to emerging areas in toxicity testing and non-testing methods. The underlying principles have not changed, but the scope of the framework has been extended to integrate information at different levels of biological organization and to reflect evolving experience in a much broader range of potential applications. These applications are relevant not only to full risk assessment for individual chemicals, but also to evolving methods for priority setting and assessment to meet increasing demands to more efficiently and accurately assess and manage large numbers of substances. They include read-across and assessment of groups of chemicals and combined exposures. The mode of action/species concordance analysis also informs hypothesis-based data generation and research priorities in support of risk assessment, related not only to (adverse) effects but also to therapeutic intervention strategies.

Envisaged broader application is illustrated in an integrative and iterative roadmap to address needs for assessment identified in formal problem formulation, as a basis to tailor the appropriate extent of mode of action/species concordance analysis. The roadmap, problem formulation and framework are iterative in nature, with feedback loops encouraging continuous refinement of fit for purpose testing strategies and risk assessment.

The relationship between mode of action and the more recently defined "adverse outcome pathway" is also clarified: conceptually, the terms are synonymous, with both representing division of the path between exposure and effect into a series of key events (including early molecular initiating events) for both individuals and populations. However, mode of action does

not necessarily imply adversity of effect, as is seemingly implied by the descriptor adverse outcome pathway.

Broader application of the modified mode of action framework is considered in two contexts, including one for which it was originally developed, where the toxicological effects of chemical exposure are known (i.e., when, as a result of problem formulation, there is a desire to perform a mode of action/species concordance analysis for an observed toxicological effect). The outcome of mode of action analysis in this application is acceptance or rejection of a hypothesized mode of action or recommendation for additional targeted research. Various case examples included here illustrate the nature of information required to demonstrate lack of human concordance, the implications of kinetic and dynamic data considered in mode of action analysis for subsequent dose–response analysis and for the design of targeted research studies using new methods (e.g., genomic technologies) and the integration of toxicological and epidemiological data.

The modified framework can also be applied in hypothesizing effects resulting from exposure to a chemical – that is, with information on putative key events in established modes of action from appropriate *in vitro* or *in silico* systems and other lines of evidence to predict and assess the likelihood of a potential mode of action and consequent effects. With the increasing amount of data available from evolving technologies, such as high-throughput and high-content screening assays, QSARs and other computational approaches, it is likely that this latter application of the framework will be of increasing value to the risk assessment community. The considerable experience acquired in the application of the framework in addressing documented (adverse) effects has a meaningful implication to inform the more limited knowledge base in these more predictive applications. This is illustrated in various case examples, including the use of mode of action analysis in prioritizing substances for further testing, in guiding development of more efficient testing strategies and in identifying critical data needs and testing strategies in read-across. In this vein, mode of action considerations should inform further development of research strategies and data generation methods, as well as the development of biomarkers.

The modified Bradford Hill considerations incorporated in framework analysis from its inception are considered a critical element to document, transparently and consistently, weight of evidence for hypothesized modes of action. These considerations have been updated and additionally articulated somewhat here to reflect increasing experience in application for cases where the toxicological outcome of chemical exposure is known. Additional work is also under way to further simplify and delineate application of the modified Bradford Hill considerations in mode of action analysis. This includes additional articulation of the modified Bradford Hill considerations for weight of evidence as a basis to contribute to common understanding, rank ordering of their importance as well as provision of examples of what might constitute strong versus weak evidence for each, based on acquired experience in mode of action analysis (Meek ME, Palermo CM, Bachman AM, North CM, Lewis RJ, submitted).

A template for extension of the concordance table in the original framework to dose–response analysis is also included, as is one for comparative consideration of weight of evidence for various modes of action based on the modified Bradford Hill considerations. Clear and transparent documentation of uncertainties at each stage of the mode of action analysis is also emphasized, with the objective of being as quantitative as possible regarding the likelihood of a hypothesized mode of action being operative in humans. Additional work to delineate more specifically the appropriate form and content of uncertainty analysis is strongly recommended, consistent with objectives and content of ongoing initiatives in this area.

Experience in mode of action analyses for documented (adverse) effects in human health risk assessment is informative in consideration of weight of evidence for hypothesized effects (referenced as adverse outcome pathways by OECD, 2012), based on early key or molecular initiating events. Based on this experience, development of proof of concept for application of the modified Bradford Hill considerations in more predictive application is strongly recommended. This is particularly important, in view of their significant reliance on demonstration of the essentiality of key events and concordance of dose–response relationships and temporality between early and late key events, information that is often lacking in the more predictive application that is envisaged. Additional collaboration between the health risk and ecological communities in this context is also recommended as a basis to draw on collective experience to increase common understanding and to develop communication and uptake strategies.

In conclusion, the modified framework and accompanying roadmap and case examples are expected to contribute to improving transparency in explicitly addressing weight of evidence considerations in mode of action and species concordance analyses based on both conventional data sources and evolving methods. The broader application envisaged here emphasizes the importance of interaction among the risk assessment, risk management and research communities, as a basis to transition to consideration of data from different levels of biological organization in fit for purpose mode of action analysis (e.g., prioritization vs. full assessment), while also highlighting the need to anchor data from evolving technologies and research. Development of the modified mode of action framework has also highlighted the conceptually identical mode of action and adverse outcome pathway and the resulting need for the research and environmental and human health risk assessment communities to move forward together to develop rigorous, efficient and transparent methodologies to meet increasingly progressive mandates to test and assess, more efficiently and more effectively, much larger numbers of chemical substances in commerce.

Acknowledgments

The authors gratefully acknowledge the European Commission Joint Research Centre for providing for a drafting group meeting and the United States Environmental Protection Agency for its support of the project. The authors also thank Kathy Hughes for contributing to the discussion and conclusions section and acknowledge Marla Sheffer for technical editing of the manuscript.

Conflict of Interest

The authors did not report any conflicts of interest.

References

Ankley GT, Bennett RS, Erickson RJ, Hoff DJ, Hornung MW, Johnson RD, Mount DR, Nichols JW, Russom CL, Schmieder PK, Serrrano JA, Tietge JE, Villeneuve DL. 2010. Adverse outcome pathways: a conceptual framework to support ecotoxicology research and risk assessment. *Environ. Toxicol. Chem.* **29**: 730–741.

Boobis AR, Cohen SM, Dellarco V, McGregor D, Meek ME, Vickers C, Willcocks D, Farland W. 2006. IPCS framework for analyzing the relevance of a cancer mode of action for humans. *Crit. Rev. Toxicol.* **36**: 781–792.

Boobis AR, Doe JE, Heinrich-Hirsch B, Meek ME, Munn S, Ruchirawat M, Schlatter J, Seed J, Vickers C. 2008. IPCS framework for analyzing the relevance of a noncancer mode of action for humans. *Crit. Rev. Toxicol.* **38**: 87–96.

Boobis AR, Daston GF, Preston RJ, Olin SS. 2009. Application of key events analysis to chemical carcinogens and noncarcinogens. *Crit. Rev. Food Sci. Nutr.* **49**: 690–707.

Cao Z, Shafer TJ, Crofton KM, Gennings C, Murray TF. 2011a. Additivity of pyrethroid actions on sodium influx in cerebrocortical neurons in primary culture. *Environ. Health Perspect.* **119**: 1239–1246.

Cao Z, Shafer, TJ, Murray TF. 2011b. Mechanisms of pyrethroid insecticide-induced stimulation of calcium influx in neocortical neurons. *J. Pharmacol. Exp. Ther.* **336**: 197–205.

Carmichael N, Bausen M, Boobis AR, Cohen SM, Embry M, Fruijtier-Pölloth C, Greim H, Lewis R, Meek ME, Mellor H, Vickers C, Doe J. 2011. Using mode of action information to improve regulatory decision-making: an ECETOC/ILSI RF/HESI workshop overview. *Crit. Rev. Toxicol.* **41**: 175–186.

Clark JM, Symington SB. 2012. Advances in the mode of action of pyrethroids. *Top. Curr. Chem.* **314**: 49–72.

Cohen SM, Boobis AR, Meek ME, Preston RJ, McGregor D. 2006a. 4-Aminobiphenyl and DNA reactivity: case study within the context of the 2006 IPCS human relevance framework for analysis of a cancer mode of action for humans. *Crit. Rev. Toxicol.* **36**: 803–819.

Cohen SM, Arnold LL, Eldan M, Lewis AS, Beck BD. 2006b. Methylated arsenicals: the implications of metabolism and carcinogenicity studies in rodents to human risk assessment. *Crit. Rev. Toxicol.* **36**: 99–133.

Cohen SM, Ohnishi T, Arnold LL, Le SC. 2007. Arsenic-induced bladder cancer in an animal model. *Toxicol. Appl. Pharmacol.* **222**: 258–263.

Council of Labor Affairs, Taiwan. 2012. Existing Chemical Substance Inventory. Chemical Substance Nomination & Notification, Council of Labor Affairs, Executive Yuan, Taiwan. Available at: http://csnn.cla.gov.tw/content/englishHome.aspx (accessed 28 September 2012).

Dellarco VL, McGregor D, Berry SC, Cohen SM, Boobis AR. 2006. Thiazopyr and thyroid disruption: case study within the context of the 2006 IPCS human relevance framework for analysis of a cancer mode of action. *Crit. Rev. Toxicol.* **36**: 793–801.

Dellarco V, Henry T, Sayre P, Seed J, Bradbury S. 2010. Meeting the common needs of a more effective and efficient testing and assessment paradigm for chemical risk management. *J. Toxicol. Environ. Health B Crit. Rev.* **13**(2–4): 347–360.

European Chemicals Bureau. 2004. European Union Risk Assessment Report: Aniline. Office for Official Publications of the European Communities, Luxembourg. Available at: http://esis.jrc.ec.europa.eu/doc/risk_assessment/REPORT/anilinereport049.pdf (accessed 28 September 2012).

European Commission. 2006. Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC,93/67/EEC, 93/105/EC and 2000/21/EC. Off. J. Eur. Union L396, 1–849. Available at: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:396:0001:0849:EN:PDF (accessed 28 September 2012).

Gocht T, Schwarz M, Berggren E, Whelan M. 2013. SEURAT-1: Development of a Research Strategy for the Replacement of *In Vivo* Repeated Dose Systemic Toxicity Testing. Available at: http://www.alttox.org/ttrc/eu/way-forward/gocht-schwarz-berggren-whelan/#1 (accessed 13 September 2013).

Hill AB. 1965. The environment and disease: association or causation? *Proc. R. Soc. Med.* **58**: 295–300.

Hou T, Li Y, Zhang W, Wang J. 2009. Recent developments of *in silico* predictions of intestinal absorption and oral bioavailability. *Comb. Chem. High Throughput Screen.* **12**: 497–506.

Hughes K, Paterson J, Meek ME. 2009. Tools for the prioritization of substances on the Domestic Substances List in Canada on the basis of hazard. *Regul. Toxicol. Pharmacol.* **55**: 382–393.

Julien E, Boobis AR, Olin SS, The ILSI Research Foundation Threshold Working Group. 2009. The key events dose–response framework: a cross-disciplinary mode-of-action based approach to examining dose–response and thresholds. *Crit. Rev. Food Sci. Nutr.* **49**: 682–689.

Knaak JB, Dary CC, Zhang X, Gerlach RW, Tornero-Velez R, Chang DT, Goldsmith R, Blancato JN. 2012. Parameters for pyrethroid insecticide QSAR and PBPK/PD models for human risk assessment. *Rev. Environ. Contam. Toxicol.* **219**: 1–114.

Lowell Center for Sustainable Production. 2012. Chemicals Policy: Other Nations [Internet]. Chemicals Policy & Science Initiative, Lowell Center for Sustainable Production, University of Massachusetts Lowell, Lowell, Massachusetts. Available at: http://www.chemicalspolicy.org/chemicalspolicy.othernations.php (accessed 28 September 2012).

McConnell ER, McClain MA, Ross J, Lefew WR, Shafer TJ. 2012. Evaluation of multi-well microelectrode arrays for neurotoxicity screening using a chemical training set. *Neurotoxicology* **33**: 1048–1057.

McGregor D, Bolt H, Cogliano V, Richter-Reichhelm HB. 2006. Formaldehyde and glutaraldehyde and nasal cytotoxicity: case study within the context of the 2006 IPCS human framework for the analysis of a cancer mode of action for humans. *Crit. Rev. Toxicol.* **36**: 821–835.

Meek ME. 2008. Recent developments in frameworks to consider human relevance of hypothesized modes of action for tumours in animals. *Environ. Mol. Mutagen.* **49**: 110–116.

Meek ME. 2009. Mode of Action Frameworks in Toxicity Testing and Chemical Risk Assessment. PhD dissertation, Institute for Risk Assessment Sciences, Utrecht University, the Netherlands.

Meek ME, Armstrong VC. 2007. The assessment and management of industrial chemicals in Canada. In *Risk Assessment of Chemicals: An Introduction*, Van Leeuwen K, Vermeire T (eds). 2nd edn. Kluwer Academic Publishers: Dordrecht, the Netherlands; 591–621.

Meek ME, Klaunig JE. 2010. Proposed mode of action of benzene-induced leukemia: interpreting available data and identifying critical data gaps for risk assessment. *Chem. Biol. Interact.* **184**: 279–285.

Meek ME, Bucher JR, Cohen SM, Dellarco V, Hill RN, Lehman-McKeeman LD, Longfellow DG, Pastoor T, Seed J, Patton DE. 2003. A framework for human relevance analysis of information on carcinogenic modes of action. *Crit. Rev. Toxicol.* **33**: 591–653.

Meek ME, Berry C, Boobis AR, Cohen SM, Hartley M, Munn S, Olin S, Schlatter V, Vickers C. 2008. Letter to the editor re: Guyton, Kathryn Z, Barone, Stanley, Jr, Brown, Rebecca C, Euling, Susan Y, Jinot, Jennifer, Makris, Susan (2008). Mode of action frameworks: a critical analysis. *J. Toxicol. Environ. Health B* **11**: 681–685.

Meek ME, Boobis AR, Crofton KM, Heinemeyer G, Raaij MV, Vickers C. 2011. Risk assessment of combined exposure to multiple chemicals: a WHO/IPCS framework. *Regul. Toxicol. Pharmacol.* **60**(2): S1–S14.

Meek ME, Palermo CM, Bachman AM, North CM, Lewis RJ. n.d. Submitted. Mode of action human relevance (MOA/HR) framework – Evolution of the modified Bradford Hill considerations and comparative analysis of weight of evidence. *Appl. Toxicol.*

[NRC] National Research Council. 2007. *Toxicity Testing in the 21st Century: A Vision and a Strategy*. National Academy Press: Washington, DC.

[OECD] Organisation for Economic Co-operation and Development. 2009. Report of the Expert Consultation to Evaluate an Estrogen Receptor Binding Affinity Model for Hazard Identification. Environment Directorate, OECD, Paris. Available at: http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=env/jm/mono(2009)33&doclanguage=en (accessed 28 September 2012).

[OECD] Organisation for Economic Co-operation and Development. 2012. Proposal for a Template, and Guidance on Developing and Assessing the Completeness of Adverse Outcome Pathways. OECD, Paris. Available at: http://www.oecd.org/chemicalsafety/testingofchemicals/49963554.pdf (accessed 28 September 2012).

Schmieder P, Mekenyan O, Bradbury S, Veith G. 2003. QSAR prioritization of chemical inventories for endocrine disruptor testing. *Pure Appl. Chem.* **75**: 2389–2396.

Schmieder PK, Tapper MA, Denny JS, Kolanczyk RC, Sheedy BR, Henry TR, Veith GD. 2004. Use of trout liver slices to enhance mechanistic

interpretation of estrogen receptor binding for cost-effective prioritization of chemicals within large inventories. *Environ. Sci. Technol.* **38**(23): 6333–6342.

Scollon EJ, Starr JM, Godin SJ, DeVito MJ, Hughes MF. 2009. *In vitro* metabolism of pyrethroid pesticides by rat and human hepatic microsomes and cytochrome p450 isoforms. *Drug Metab. Dispos.* **37**: 221–228.

Seed J, Carney EW, Corley RA, Crofton KM, DeSesso JM, Foster PM, Kavlock R, Kimmel G, Klaunig J, Meek ME, Preston RJ, Slikker W, Jr, Tabacova S, Williams GM, Wiltse J, Zoeller RT, Fenner-Crisp P, Patton DE. 2005. Overview: Using mode of action and life stage information to evaluate the human relevance of animal toxicity data. *Crit. Rev. Toxicol.* **35**: 664–672.

Soderlund DM. 2012. Molecular mechanisms of pyrethroid insecticide neurotoxicity: recent advances. *Arch. Toxicol.* **86**: 165–181.

Sonich-Mullin C, Fielder R, Wiltse J, Baetcke K, Dempsey J, Fenner-Crisp P, Grant D, Hartley M, Knaap A, Kroese D, Mangelsdorf I, Meek E, Rice JM, Younes M, International Programme on Chemical Safety. 2001. IPCS conceptual framework for evaluating a mode of action for chemical carcinogenesis. *Regul. Toxicol. Pharmacol.* **34**: 146–152.

[US EPA] US Environmental Protection Agency. 1996. Proposed guidelines for carcinogen risk assessment. *Fed. Reg.* **61**: 17960–18011.

[US EPA] US Environmental Protection Agency. 2005a. Guidelines for Carcinogen Risk Assessment. US EPA, Washington, DC. Available at: http://www.epa.gov/cancerguidelines/

[US EPA] US Environmental Protection Agency. 2005b. Science Issue Paper: Mode of Carcinogenic Action for Cacodylic Acid (Dimethylarsinic Acid, DMA^V) and Recommendations for Dose Response Extrapolation. Health Effects Division, Office of Pesticide Programs, US EPA, Washington, DC.

[US EPA] US Environmental Protection Agency. 2009a. An Effects-based Expert System to Predict Estrogen Receptor Binding Affinity for Food Use Inert Ingredients and Antimicrobial Pesticides: Application in a Prioritization Scheme for Endocrine Disrupting Screening [Draft]. Office of Pesticide Programs, US EPA, Washington, DC. Available at: http://www.regulations.gov/#!documentDetail;D = EPA-HQ-OPP-2009-0322-0002 (accessed 28 September 2009).

[US EPA] US Environmental Protection Agency. 2009b. Integrated Science Assessment for Particulate Matter (Final Report). EPA/600/R-08/139F. US EPA, Washington, DC, pp. 5-1–5-23.

Review Article

Journal of Applied Toxicology

Received: 2 November 2013, Revised: 28 November 2013, Accepted: 3 December 2013 Published online in Wiley Online Library

(wileyonlinelibrary.com) DOI 10.1002/jat.2984

Mode of action human relevance (species concordance) framework: Evolution of the Bradford Hill considerations and comparative analysis of weight of evidence

M. E. (Bette) Meek*, Christine M. Palermo, Ammie N. Bachman, Colin M. North and R. Jeffrey Lewis

ABSTRACT: The mode of action human relevance (MOA/HR) framework increases transparency in systematically considering data on MOA for end (adverse) effects and their relevance to humans. This framework continues to evolve as experience increases in its application. Though the MOA/HR framework is not designed to address the question of "how much information is enough" to support a hypothesized MOA in animals or its relevance to humans, its organizing construct has potential value in considering relative weight of evidence (WOE) among different cases and hypothesized MOA(s). This context is explored based on MOA analyses in published assessments to illustrate the relative extent of supporting data and their implications for dose–response analysis and involved comparisons for chemical assessments on trichloropropane, and carbon tetrachloride with several hypothesized MOA(s) for cancer. The WOE for each hypothesized MOA was summarized in narrative tables based on comparison and contrast of the extent and nature of the supporting database versus potentially inconsistent or missing information. The comparison was based on evolved Bradford Hill considerations rank ordered to reflect their relative contribution to WOE determinations of MOA taking into account increasing experience in their application internationally. This clarification of considerations for WOE determinations as a basis for comparative analysis is anticipated to contribute to increasing consistency in the application of MOA/HR analysis and potentially, transparency in separating science judgment from public policy considerations in regulatory risk assessment. Copyright © 2014. The Authors. Journal of Applied Toxicology Published by John Wiley & Sons Ltd.

Keywords: human relevance framework; mode of action; weight of evidence; key events; evolved Bradford Hill considerations

Introduction

The mode of action/human relevance (MOA/HR) framework is an analytical framework designed to increase transparency in the systematic consideration of the weight of evidence (WOE) of hypothesized MOA(s) for critical effects and their relevance to humans. It was developed in initiatives of the International Life Sciences Institute Risk Sciences Institute (ILSI RSI) and the International Programme on Chemical Safety (IPCS) and derives from earlier work on MOA by the US Environmental Protection Agency (USEPA) and IPCS (Sonich-Mullin *et al.*, 2001).

The development and evolution of the IPCS ILSI RSI MOA/HR framework, which has involved large numbers of scientists internationally, is described in several publications (Boobis *et al.*, 2006, 2008; Meek, 2008; Meek *et al.*, 2003; Seed *et al.*, 2005). Potential application in a broader range of relevant contexts has been considered more recently (Carmichael *et al.*, 2011; Meek and Klaunig, 2010). The framework has been illustrated by an increasing number of case studies ($n = 30$, currently), and is widely adopted in international and national guidance and assessments (Meek *et al.*, 2008), including those of the USEPA (Dellarco and Baetcke, 2005; Manibusan *et al.*, 2007; SAB, 1999, 2007; SAP, 2000; USEPA, 2005a). Building on this collective experience, the framework has been updated recently, to address uncertainty additionally and to extend its utility to emerging areas in toxicity testing and non-testing methods. The update includes incorporation within a roadmap, encouraging continuous refinement of fit-for-purpose testing strategies and risk assessment (Meek *et al.*, 2014).

In addition to increasing transparency through structured articulation of the evidence and uncertainties upon which conclusions are based, MOA/HR analysis also contributes to the transparent assimilation of all available data in both a risk assessment and research context. This is important because it facilitates identification of critical data needs and contributes to transparency in the separation of science judgment (i.e., weighting of options based on systematic consideration of available scientific support) from public health protection policy, the latter

*Correspondence to: M. E. (Bette) Meek, University of Ottawa, One Stewart Street, Suite 309, Ottawa, Ontario, Canada, K1N 6N5.
Email: bmeek@uottawa.ca

University of Ottawa, One Stewart Street, Suite 309, Ottawa, Ontario, K1N 6N5, Canada

sometimes involving embedded conservatism to increase public health protection.

Though the MOA/HR framework is not designed to address the question of "how much information is enough" to support a hypothesized MOA in animals or its relevance to humans, its organizing construct has value in considering relative WOE among different cases and hypothesized MOAs. Comparative WOE evaluation for MOA/HR analysis is illustrated as a basis to increase common understanding of the nature of transparency required to document the relative degree of confidence in supporting data for hypothesized MOAs. To demonstrate this approach, WOE for MOA/HR analysis in two published assessments (i.e., carbon tetrachloride and 1,2,3-trichloropropane [TCP]) (USEPA, 2009, 2010) is comparatively considered in the context of evolved Bradford Hill (B/H) considerations introduced here to promote better common understanding and consistency in use. The focus here is not on the conclusions of the assessments but rather, the utility of comparative analysis for WOE evaluation in MOA/HR analysis. These cases were specifically selected to exemplify varying degrees of WOE for several hypothesized MOA.

Methods And Results

Details of the updated MOA/HR framework are available elsewhere (Meek *et al.*, 2014). Briefly, the WOE for a hypothesized MOA in animals is assessed based on considerations modified from those proposed by Bradford Hill (Hill, 1965) for assessment of causality in epidemiological studies. HR or species concordance is then systematically considered, taking into account more generic information such as anatomical, physiological and biochemical variations. If the WOE for the hypothesized MOA is sufficient and relevant to humans, implications for dose–response in humans are then considered in the context of kinetic and dynamic data. Delineation of the degree of confidence in the WOE for hypothesized MOAs is critical, as is the delination of critical research needs.

Establishing support for or rejection of a hypothesized MOA provides the foundation for subsequent considerations of dose–response, HR and estimates of risk. It involves (1) delineation of key events leading to the end (adverse) effect in a hypothesized MOA and (2) evaluation of all of the data to consider the extent of the supporting WOE for the hypothesized MOA. Importantly, if alternative MOA(s) are supported, these are evaluated with equal rigor in separate MOA/HR framework analyses. Ultimately, depending upon the application, there may be a need to draw a conclusion on the sufficiency of data supporting a MOA, to assess different risk management options. The comparative analysis of WOE was developed as a basis for increasing common understanding of the nature of transparency required to document the degree of confidence in the sufficiency of supporting data for hypothesized (potentially competing) MOAs.

A template for WOE analysis of MOA based on the evolved B/H considerations is presented in Table 1. In this approach, supporting data, inconsistent data and missing information are evaluated and tabulated in the context of the evolved B/H considerations presented here. The data in this table are considered in totality to assess the WOE for a MOA. In addition, the evidence can be used in a comparative manner to gain perspective on the relative degree of confidence that a hypothesized MOA is operative, based on the extent of supporting WOE compared to that for another postulated MOA for the same chemical or for the same MOA for other chemicals.

As illustrated in Table 1, WOE analysis is heavily dependent on the B/H considerations. Previous iterations of modified B/H considerations have been applied inconsistently in MOA/HR analyses, which may be attributable in large measure to the availability of only relatively general, early guidance in this area (USEPA, 2005b; Sonich-Mullin *et al.*, 2001). Some of the considerations have been misinterpreted due to a lack of common understanding of their appropriate level of application to MOA data in a WOE context; i.e., in overall data synthesis and evaluation of sufficiency of evidence to support a MOA decision versus the initial phase of systematic review (i.e., data selection and individual study review). Table 2 summarizes the variation in definitions of the B/H considerations in MOA analysis, which may also have contributed to inconsistency in application.

Evolved B/H considerations have been proposed and clarified here through delineation of the specific aspects addressed by each, as framed by a series of questions (captured below and summarized in Table 3). These questions build on those presented in Meek *et al.* (2014), based on additional experience in considering transparency in existing assessments as a basis to document comparative WOE. These evolved B/H considerations are proposed, then, not only as a basis to increase consistency in making WOE determinations for hypothesized MOA(s), but also to

Table 1. Template for weight of evidence based on evolved Bradford Hill considerations

Evolved Bradford Hill Considerations		Supporting Data	Inconsistent Data	Missing Data
1. Biological Concordance				
2. Essentiality of Key events				
3. Concordance of Empirical Observations among Key Events	Dose–response Temporality Incidence			
4. Consistency				
5. Analogy				

For a postulated mode of action, supporting data, inconsistent data and missing data are tabulated in the context of the evolved Bradford Hill considerations. Input in the supporting and inconsistent columns captures only what has been observed. Input in the missing column includes only that which is technically feasible and that is important for informing the mode of action. Cells are left blank in instances where data do not exist or are inadequate for evaluation. A brief narrative should accompany this table to describe the overall determination as to whether the data support or refute the hypothesis.

Table 2. Definition of the Bradford Hill considerations for application in mode of action analysis

Bradford Hill Considerations (Hill, 1965)	IPCS MOA/HR Framework (Boobis *et al.*, 2006; 2008; Sonich-Mullin *et al.*, 2001)	EPA Cancer Guidelines (USEPA, 2005b)	Evolved Bradford Hill considerations
Strength Strength of the association between suspected cause and observation.	**Strength** Unclearly defined. Considered together with specificity and consistency.	**Strength** The finding of large risks increases confi dence the association is not due to chance.	**N/A** Not considered applicable for evaluating MOA data.
Consistency Repeatability of an association by different persons, in different places, circumstances and times.	**Consistency** Repeatability of the key events in different studies. Considered together with strength and specificity.	**Consistency** Pattern of elevated risk observed across several independent studies.	**Consistency** Is the pattern of effects across species/ strains/ organs/test systems what would be expected?
Specificity The association is limited to a specific population and to particular sites and types of disease.	**Specificity** Stop/recovery studies show an absence or reduction of toxicity when a key event is blocked or reduced. Considered together with strength and consistency.	**Specificity** One cause associated with a single effect or disease.	**Essentiality of key events** Is the sequence of events reversible if dosing is stopped or a key event prevented?
Temporality The exposure occurs before the effect.	**Temporal association** Key events should be observable before toxicity is apparent.	**Temporal relationship** When exposure is known to precede development of the disease.	**Temporal concordance** Are the key events observed in hypothesized order?
Biological gradient Risk of disease increases with increasing exposure.	**Dose–response relationship** The dose–response for key events parallel the dose–response for the toxic effect. Increases in incidence of a key event correlate with increase in incidence of later key events.	**Biological gradient** Increasing effects associated with greater exposure.	**Dose–response concordance** Are the key events observed at doses below or similar to those associated with the end (adverse) effect?
Plausibility Biological knowledge supports suspected causation.	**Biological plausibility and coherence** Consistent with current understanding of biology. Considered together with coherence.	**Biological plausibility** Consistency with data from experiments or other sources demonstrating biological plausibility.	**Biological concordance** Does the hypothesized MOA conflict with broader biological knowledge? How well established is the MOA in the wider biological database?
Coherence The association agrees with the generally known facts of the history and biology of the disease.	**Coherence** Consistency with what is known specifically about the overall biological effects of the substance. Considered together with biological plausibility.	**Coherence** Information supporting cause and effect from other lines of evidence (i.e., animal bioassays, toxicokinetic studies and short-term studies).	**N/A** Not considered applicable for evaluating MOA data
Experiment Experimental evidence alters the frequency of associated events.	**N/A** Has not been mentioned in recent publications on the MOA/HR framework.	**Experimental evidence** when a change of exposure in a human popu lation brings about a change in disease.	**N/A** Not considered applicable for evaluating MOA data.

(Continues)

Table 2. *(Continued)*

Bradford Hill Considerations (Hill, 1965)	IPCS MOA/HR Framework (Boobis *et al.*, 2006; 2008; Sonich-Mullin *et al.*, 2001)	EPA Cancer Guidelines (USEPA, 2005b)	Evolved Bradford Hill considerations
Analogy Information for a similar but different association supports causation. **N/A**	**N/A** Has not been mentioned in recent publications on the HR/MOA framework. **N/A** Considered as part of dose–response relationship definition.	**Analogy** insight gained from structure activity relationships and information on structural analogues. **N/A**	**Analogy** Would the MOA be anticipated based on broader chemical specific knowledge? **Incidence concordance** Is the occurrence of the end (adverse) effect less than that for preceding key events?

HR, human relevance; MOA, mode of action.

promote consistency in their application based on accumulating experience internationally.

The evolved B/H considerations are described in more detail below. These considerations appear in rank order based on their appropriate weighting of relative contribution to WOE determinations for hypothesized MOA(s), with those listed first contributing most significantly. Examples for evaluating weak to strong evidence for each evolved B/H consideration are also discussed.

Biological Concordance

- Does the hypothesized MOA conflict with broader biological knowledge?
- How well established is the MOA?

Evidence for a hypothesized MOA must satisfy the consideration of biological concordance. If available data on the hypothesized MOA are at odds with biological understanding, the hypothesis does not constitute a reasonable option for consideration. For instance, if a hypothesized early key event cannot conceivably lead to a subsequent hypothesized key or end event, it need not be considered.

The extent of evidence for biological concordance would be considered stronger, for example, if the hypothesized MOA has been well documented for a broad range of chemicals, and weaker if the hypothesized MOA is conceivable based on limited data or it has been hypothesized based simply on the possibility that none of the key events are at odds with biological understanding.

Essentiality of Key Events

- Is the sequence of events reversible if dosing is stopped or a key event prevented (i.e., counterfactual evidence)?

The extent of counterfactual evidence (i.e., experimental support for the necessity of a key event) is one of the principal determinants of WOE for a hypothesized MOA (Borgert *et al.*, 2011). For example, experimental evidence in animal models that lack a key metabolic pathway (e.g., knock out animal models) and fail to develop the end (adverse) effect would support essentiality of a key event. Similarly, if following cessation of repeated exposure for various periods, effects are reversible (i.e., late key events and/or the end (adverse) effect is prevented), this constitutes relatively strong evidence that key events are causal.

It is important to note that by its nature, counterfactual evidence typically addresses the necessity of an individual key event in a hypothesized MOA. Therefore, it may not always be helpful for discerning between two possible MOAs that share a key event. For example, if a chemical requires metabolic activation to be carcinogenic, a negative result in a 2-year cancer bioassay in an animal model null for the necessary activating enzyme supports that metabolism is necessary for carcinogenesis but is not helpful for differentiating between a MOA involving metabolic activation followed by direct DNA damage versus a MOA involving metabolic activation followed by cytotoxicity and regenerative proliferation.

Support for the essentiality of key events is considered stronger when there is direct counterfactual evidence supporting multiple key events in the hypothesized MOA. Evidence is considered weaker when evidence involves indirect measures for key events (i.e., the key event is inferred from the actual measured endpoint)

Journal of Applied Toxicology

Table 3. Proposed changes to the Bradford Hill considerations and guidance for interpretation to improve application in the MOA/HR framework[a]

Evolved Bradford Hill considerations	Defining questions	Evidence for evaluating degree of support for the mode of action	
		Stronger	Weaker
1. Biological Concordance (replaces biological plausibility & coherence)	Does the hypothesized MOA conflict with broader biological knowledge? How well established is the MOA?	MOA is well established in scientific knowledge and/or completely consistent with established biological understanding.	MOA is contrary to well established biological understanding. MOA requires biological processes that are novel or poorly established.
2. Essentiality of Key Events (replaces strength, and specificity)	Is the sequence of events reversible if dosing is stopped or a key event prevented?	Counterfactual evidence to support key events (e.g., absence/reduction of later events when an earlier key event is blocked or diminished).	Data on reversibility only, indirect evidence only for key events or limited data available to assess.
3. Concordance of Empirical Observations among Key events (encompasses dose response and temporal concordance and beyond)	Dose–response: Are the key events observed at doses below or similar to those associated with end (adverse) effect? Temporality: Are the key events observed in hypothesized order? Incidence: Is the occurrence of the end (adverse) effect less than that for the preceding key events?	Dose–response and temporality: expected pattern of temporal and dose–response relationships based on robust database (multiple studies with examination of key events at interim time periods and at least 3 doses). Incidence: incidence of early key events is greater than end (adverse) effect.	All key events occur at all dose levels and all time points and/or limited data available to assess (e.g., inadequate dose spacing, missing key time periods for effect development, or failure to assess incidence at early time points). Incidence of early key events is lower than the end (adverse) effect and/or limited data available to assess.
4. Consistency (among different biological contexts)	Is the pattern of observations across species/strains/organs/test systems what would be expected based on the hypothesized MOA?	Pattern of effects are what would be expected across species, strains, organs and/or test systems.	Significantly inconsistent pattern of effects or limited data available to assess (e.g., effect only observed in a single rat strain).
5. Analogy (consistency across chemicals)	Would the MOA be anticipated based on broader chemical specific knowledge (e.g., the chemical is a member of a category for which related chemicals have known or strongly suspected MOA)?	Observations are consistent with those for other (related) chemicals having well defined MOA.	Pattern of effects for other (related) chemicals is distinctly different. Insufficient data to evaluate whether chemical behaves like related chemicals with similar proposed MOA.

MOA, mode of action.

[a]Evolution of the Bradford Hill (B/H) considerations for improved fit-for-purpose in the evaluation of sufficiency of data to support a hypothesized MOA. The evolved B/H considerations are rank ordered based on their appropriate weighting of relative contribution to weight of evidence determinations for hypothesized MOA(s), with those listed at the top contributing most significantly.

or non-specific inhibition of key events. For example, for a MOA hypothesized to involve binding to a receptor, demonstrating an end (adverse) effect is prevented by knocking-out or downregulating expression of the receptor is stronger than counterfactual evidence using a non-specific inhibitor.

Concordance of Empirical Observation Among Key Events

Concordance of empirical observations contributes considerably to the WOE for hypothesized MOA(s). Specifically, concordance of dose–response, temporality and incidence are key considerations. Each of these is addressed separately below. While not weighted as heavily as biological concordance and essentiality of key events, concordance of empirical observation across dose–response, temporality and incidence contributes significantly to WOE. Relationships and outliers should be carefully evaluated to understand whether the WOE strongly supports or is discordant with the hypothesized MOA, including consideration of cohesiveness across all three aspects of empirical observation.

Concordance of Dose–response Relationships Among Key Events

- Are the key events observed at doses below or similar to those associated with the end (adverse) effect?

In past MOA analyses, assessment of dose–response has sometimes been misinterpreted as simply addressing the question: "Is there evidence of a dose–response relationship for key events and/or the end (adverse) effect?" While this question is relevant to hazard characterization, it does not address dose–response concordance in relation to the WOE for a hypothesized MOA. Rather, the latter addresses the consistency of observed dose–response relationships among key and end (adverse) effects, as framed explicitly in the question above.

The hypothesized MOA is not supported in scenarios for which there is evidence that early key events occur only at higher doses than the end (adverse) effect. For example, a hypothesized receptor-based MOA is not supported by evidence indicating that receptor binding occurs only at doses well above those that cause frank liver injury, though it is important to consider if this might be a function of dose spacing in the relevant studies. Benchmark dose analyses for the dose–response relationships in key and end events are the most appropriate measure for consideration of their concordance, as they provide for direct comparison of comparable doses associated with a specified increase in each of the key events and/or end (adverse) effects and normalize for variations in dose spacing and group sizes in different studies.

Examination of the pattern of dose–response relationships is particularly important in considering the degree of support for hypothesized mutagenic MOAs (i.e., where mutation is an early and influential key event). For example, observation of a mutagenic response at high (cytotoxic) doses in genotoxicity assays is supportive of hypothesized MOAs where mutation is a secondary consequence of increased proliferative response resulting from tissue damage.

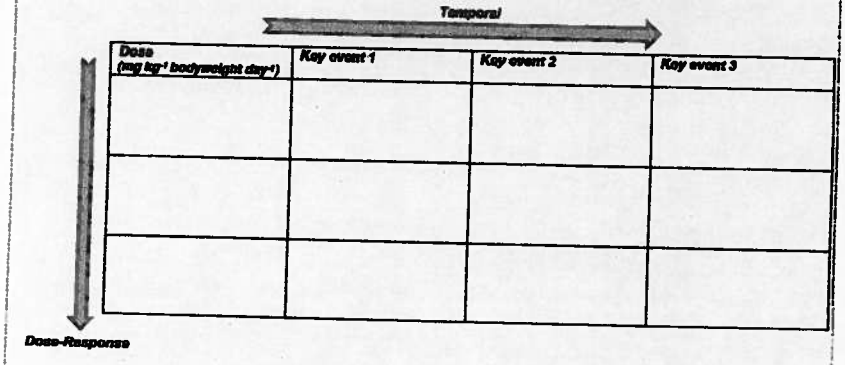
Concordance of Temporality (Time) Among Key Events

- Are the key events observed in hypothesized order?

Temporal concordance refers to the observation of key events in sequential order as described in the hypothesized MOA. In other words, earlier key events should be observed to precede later key events and the late (adverse) effect. Stronger evidence for temporal concordance is obtained when key events at interim time points demonstrate the hypothesized order (either in a single robust study or across multiple studies). Such evidence can often be acquired in studies examining the reversibility of key events and end (adverse) effects following various periods of exposure. Weaker evidence occurs when temporal data on key events are missing.

The template presented in Table 4 is often helpful in determining the extent to which evidence fulfills consideration of dose–response and temporal concordance in WOE analysis for MOA. If the hypothesized MOA is supported, the table should fill diagonally from the top left-hand corner to the bottom right-hand corner. This "pattern" supports a continuum of the relationship between early key events occurring at lower doses than late key events and outcome. Evidence of dose–response and temporal concordance is, for example, weaker if all key events occur at all dose levels and time points. Evidence is stronger, for example, if there is a reasonable range of studies of different durations with a minimum of three dose levels each and the "pattern" of results in this table (Table 4) is as described above.

Table 4. Dose–response and temporal concordance analysis template



Source: Meek and Klaunig (2010).

Journal of Applied Toxicology

Concordance of Incidence Between Key Events and End (Adverse) Effects

- Is the occurrence of the end (adverse) effect less than that for the preceding key events?

Clear evidence of the concordance of the incidence of the end (adverse) effect with that for early hypothesized key events is influential in contributing to WOE for hypothesized MOA(s). The incidence of hypothesized early key events should be greater than that for later key events and the (adverse) outcome, consistent with the important biological underpinning that key events are essential but not necessarily sufficient, to induce the relevant end (adverse) effect. For example, the hypothesis that cytotoxicity followed by regenerative proliferation are key events in the induction of specific tumors would be supported by the observation that the incidence of the former (cytotoxicity/regenerative proliferation) is greater than that for the latter (tumors) at a similar dose. "Incidence" here refers to the occurrence of a lesion of defined severity for each of the key and end events. It should be noted that a 1: 1 correlation of the incidence of early and late key events is not anticipated; lack of evidence for a 1: 1 correlation does not detract from contribution to the overall WOE. Consistent with the essentiality (but not necessarily sufficiency) of key events, lack of 1: 1 concordance is not unexpected, being a function of biological variability; i.e., lesions will not have progressed to the end (adverse) effect in all animals at the termination of exposure.

Consistency

- Is the pattern of observations across species/strains/organs/test systems what would be expected based on the hypothesized MOA?

Evidence of internal consistency within the collective data set for a chemical contributes to increased confidence in the WOE supporting a MOA. For example, if the initial hypothesized key event is oxidative metabolism to a reactive intermediate, are the target tissues and organs those which would be expected based on knowledge of distribution of the relevant metabolic enzyme? Evidence of consistency is stronger if the pattern of species-, strain- and sex-related variations in response is what would be expected based on known differences in metabolic profiles (e.g., extent and rate of metabolism to the putatively toxic entity). Evidence is weaker if there is either significant inconsistency in the expected pattern of the collective data based on the hypothesized MOA (e.g., the effect or result is only demonstrated in a single rat strain when data are available for multiple strains, for all of whom metabolizing capacity for the relevant pathway is anticipated to be similar) or when there are limited data available to assess this aspect.

Analogy

- Would the MOA be anticipated based on broader chemical specific knowledge?

Convincing evidence that the hypothesized MOA is operative for a broad range of chemically similar substances also contributes significantly to WOE. For example, consider the case where reductive metabolism for chemically similar substances is associated with a particular pattern of observations leading to the end (adverse) effect. If the pattern of observations for a related chemical is distinctly different, the evidence is weaker that these effects are produced by a similar MOA. On the other hand, if there is an extensive database illustrating that the MOA of interest is operative and leads to similar end (adverse) effects for several closely structurally related chemicals as identified, for example, by (quantitative) structure–activity modeling, evidence is stronger.

The rank order of the B/H considerations suggested above reflects their relative contribution to WOE determinations of MOA and is based on evolving experience internationally. In essence, data that conflicts with a broader biological understanding ranked highly here may be grounds for considering the available supporting data as inconsistent with the hypothesized MOA, whereas lack of concordance of some empirical data is often due to variations in, for example, dose spacing or administered doses in various studies and based on careful evaluation, would not detract meaningfully from the supporting database. In assessing the totality of the WOE, it is helpful to systematically take into account all of the considerations presented here as a basis to contribute to transparency in decision making. Such assessment benefits most from multidisciplinary input from both the relevant research and risk assessment communities. However, there is no minimum number of these evolved B/H considerations that must be met to determine sufficiency and/or associated confidence but rather, in their careful, systematic, more transparent and consistent consideration, cohesiveness (or not) of the supporting data becomes evident. It is also important to recognize that while some of the evolved B/H considerations may address the association of just one key event to the end event (e.g., essentiality of key events) the WOE determination is based on consideration of the interdependence of the key and end events in the hypothesized MOA.

Comparative Weight of Evidence Case Studies

To illustrate the utility of the comparative WOE approach, assessments for two chemicals (USEPA, 2009, 2010) were selected as case studies (i.e., carbon tetrachloride and TCP). The assessment of carbon tetrachloride drew on a previous evaluation of the US EPA (Manibusan *et al.*, 2007), though the conclusions varied. These assessments were chosen based on the condition that B/H considerations for WOE had been explicitly addressed, consistent with the analysis in the MOA/HR framework for several potential MOA(s) for carcinogenicity. The focus here was not on the conclusions of the assessments; rather, the extensive review and synthesis of data therein provided the opportunity to address the potential utility of comparative analysis based on the evolved B/H considerations for WOE in MOA/HR analysis. As such, the evidence and conclusions were not re-evaluated but were simply extracted from the referenced assessments and summarized in the narrative tables presented (Tables 5a,b and 6) for the purpose of illustrating the methodology. Similarly, assessment of the underlying investigations was not considered, though based on the approach presented here, this might constitute an important next step. The literature reviews were also not updated, as the current analysis does not focus on particular chemicals but rather the potential value of the proposed methodology.

Carbon Tetrachloride

This analysis is based on a published hazard and dose–response assessment for carbon tetrachloride (USEPA, 2010). Carbon

Table 5. (a) Comparative weight of evidence analysis for carbon tetrachloride: cytotoxic MOA[a]

Evolved Bradford Hill considerations		Supporting data	Inconsistent data	Missing data
1. Biological concordance		Sustained cytotoxicity and proliferation is a well-established MOA for chemically mediated carcinogenicity.		
2. Essentiality of key events		No carbon tetrachloride induced liver toxicity in CYP2E1 knockout mice. CYP450 inhibitors prevent carbon tetrachloride liver damage. Mice treated with CYP450 inducers have increased carbon tetrachloride toxicity in subchronic and chronic studies.		
3. Concordance of empirical observations	Dose–response	Cytotoxicty and proliferation are observed at doses equal to or lower than doses at which tumors develop in rats and male mice	Tumors elevated at the lowest dose tested in female mice (5 ppm) without hepatocellular damage.	
	Temporality	Progression from cytotoxicity to hepatocellular proliferation is supported in acute and subchronic studies in rodents. Temporal relationship of cytotoxicity, repair, proliferation and tumor development is also supported in chronic cancer bioassay in rats.		Temporal relationship in female mice is not clearly defined.
	Incidence			
4. Consistency		Hepatic toxicity, necrosis and regenerative proliferation have generally been reported in animals exposed to carbon tetrachloride orally or by inhalation and are correlated with CYP450 content. Some evidence of DNA damage observed in concert with cytotoxicity.	One study reported development of tumors in mice at doses that did not produce necrosis but design of study may have influenced this result as animals were killed 1 month after last treatment.	
5. Analogy				

MOA, mode of action.

[a]All conclusions in the above tables were extracted from the original US EPA toxicology review on carbon tetrachloride (USEPA, 2010).

(b) Comparative weight of evidence analysis for carbon tetrachloride: mutagenic MOA[a]

1. Biological concordance		Genotoxic MOA is well established for chemically mediated carcinogenicity.		
2. Essentiality of key events				
3. Concordance of empirical observations	Dose–response		Genotoxicity generally found at doses with cytotoxic effects.	Measurement of genetic damage to DNA has not been well

Journal of AppliedToxicology

		characterized at dose levels that do not cause cytotoxicity.
Temporality	Temporality not observed. Genotoxicity generally found in concert with cytotoxicity.	
Incidence	Extensive *in vitro* and *in vivo* genotoxic data are primarily negative.	Genetically damaging events occurring at or below doses that induce cytotoxicity in laboratory rodents.
4. Consistency	Doses where cytotoxic events are observed are lower than doses for which mutagenicity has been evaluated.	
	Limited positive results in genotoxicity assays appear more related to a cytotoxic response than to a mutation event	
5. Analogy		

MOA, mode of action.
[a]All conclusions in the above tables were extracted from the original US EPA toxicology review on carbon tetrachloride (USEPA, 2010).

tetrachloride caused hepatocellular adenomas and carcinomas in rats, mice and hamsters in oral studies and in rats and mice following inhalation exposure. In addition to liver tumors, adrenal pheochromocytomas were observed in male and female mice following oral and inhalation exposure, for which it was concluded that data were inadequate to evaluate MOA. There was no increase in pheochromoctyomas in rats.

Based on the analysis of available data, including that on MOA, it was concluded in the assessment (USEPA, 2010) that the agent is likely a human carcinogen. Further, a potential MOA for carbon tetrachloride-induced liver tumors was hypothesized, with the following key events that included: (1) metabolism to the trichloromethyl radical by CYP2E1 and subsequent formation of the trichloromethylperoxy radical; (2) radical-induced damage leading to hepatocellular toxicity; and (3) sustained regenerative and proliferative changes in the liver in response to hepatotoxicity. The possibility that carbon tetrachloride may act via a mutagenic MOA (i.e., where mutation is an influential early key event in the induction of tumous versus, for example, being secondary to tissue damage) was also considered but not evaluated in a manner based on WOE considerations consistent with the MOA/HR framework. Based on the inconsistencies in the database supporting a potential role for the cytotoxicity, regenerative, proliferation-based MOA at the low end of the experimental exposure range and the complexity of the genotoxicity database, it was concluded that, "... the carcinogenic MOA for carbon tetrachloride is not known. Therefore, consistent with the *Guidelines for Carcinogen Risk Assessment* (USEPA, 2005b), linear low-dose extrapolation as a default approach was applied to data for liver tumors and pheochromocytomas" (USEPA, 2010).

1,2,3-Trichloropropane

This analysis is based on a hazard and dose–response assessment of TCP released in 2009 (USEPA, 2009). Based on the observed statistically significant dose-related increases in multiple tumor types in both sexes of rats and mice in a 2-year carcinogenicity assessment (NTP, 1993) and related mechanistic data (including that on genotoxicity), it was concluded that TCP is "likely to be carcinogenic to humans" via a mutagenic MOA. Relevant data for alternative MOA(s) such as cytotoxicity with tissue repair and disruption of cell signaling were considered insufficient to evaluate. It was further concluded that the available data support a hypothesized mutagenic MOA with two key events: (1) metabolism to a DNA-reactive compound, and (2) (early) induction of mutations. A low-dose linear extrapolation approach to dose–response analysis was applied, consistent with the *Guidelines for Carcinogen Risk Assessment* (USEPA, 2005b)

Comparative Weight of Evidence Analysis

Narrative comparative WOE summary tables were constructed for the hypothesized and alternative MOA(s) for carbon tetrachloride (Table 5a,b) and for a mutagenic MOA for TCP (Table 6) based on the consideration and evaluation of the data in the existing assessments (USEPA, 2009, 2010). For each postulated MOA, supporting data, inconsistent data and missing information were tabulated in the context of the evolved B/H considerations. As per MOA/HR framework recommendations, the information in the supporting and inconsistent data columns capture what has been observed, not what might be possible if more experiments had been performed. In addition, the

Table 6. Comparative weight of evidence analysis for 1,2,3-trichloropropane: mutagenic MOA

Evolved Bradford Hill considerations		Supporting data[a]	Inconsistent data[a]	Missing data[b]
1. Biological concordance		Genotoxic MOA is well established for chemically mediated carcinogenicity		
2. Essentiality of key events		Inducers/inhibitors of metabolism alter amount of DNA binding		
3. Concordance of empirical observation	Dose–response	Dose-related formation of DNA-reactive metabolite, DNA adduct formation, tumor formation and time to tumor.		Evidence for adduct conversion to genetic damage
	Temporality	Metabolism to reactive intermediate occurs within hours of exposure, adducts appear within hours and days of exposure, and tumors first appear after ≈ 9 months.		
	Incidence			No data to assess whether adduct formation frequency different from tumor frequency.
4. Consistency		Mutagenic effects *in vitro* accompanied by limited evidence of *in vivo* mutagenicity.	Adducts occur in tissues where no neoplastic effects were reported (spleen, liver and glandular stomach). Negative results from *in vivo* genotoxicity assessments (dominant lethal and micronucleus).	
5. Analogy		Other halogenated aliphatic chemicals (1,2,-dibromoethane and 1,2-dibromo-3-chloropropane) are mutagenic carcinogens. Other genotoxic chemicals are multisite and multispecies carcinogens.		

MOA, mode of action.

[a]All conclusions in the above tables were extracted from the original US EPA toxicology review on 1,2,3-trichloropropane (USEPA, 2009).

[b]The IRIS assessment did not comment on missing data; the information here represents the authors' views.

information noted in the missing column only includes that which is testable and important for informing the MOA (i.e., critical data needs). Ideally, a discussion on whether the missing information is critical and would detract from or impact conclusions regarding the proposed MOA should accompany this comparative WOE table. Blank cells would typically represent instances where data either do not exist or are inadequate for evaluation. However, in this case, as the analysis draws upon an existing assessment, blank cells may also represent where text was either absent or inadequate to address the evolved B/H considerations.

Qualitative Assessment of Overall Evidence

For both case studies, the focus is not to conclude on the sufficiency of underlying data to support a particular MOA conclusion, but rather to illustrate the utility of the comparative WOE approach for increasing transparency in the assimilation of data.

Visually, Tables 5(a,b) and 6 highlight the availability of supporting and discrepant data on the MOA(s) evaluated for carbon tetrachloride and TCP. Comparative WOE analysis, for the two hypothesized MOA(s) for carbon tetrachloride based on the published assessment (USEPA, 2010), indicates that the supporting data for the hypothesized MOA involving cytotoxicity (necessarily within the range of experimental observation) fulfill a number of the evolved B/H considerations. This contrasts with the comparatively more limited support for the hypothesized mutagenic MOA. This difference highlights:

(1) the potential utility of comparative analysis for assessing the WOE of alternative MOA(s) for individual chemicals, based on the evolved B/H considerations to more explicitly indicate the degree of confidence in a particular MOA, and
(2) the desirability, in the interest of transparency and consistency, of separating conclusions reflecting assessment of the relative WOE for MOA in the observable experimental range based on articulated and explicit considerations from those based on inference or extrapolation to the low-dose range. It is anticipated that such an approach has the potential to increase transparency in delineating science judgment determinations from those related to public policy.

The comparative WOE analysis for TCP also provides a basis for comparison across chemicals of a relatively strong database for a mutagenic MOA, which can be contrasted with one that is relatively weak, potentially as a basis to increase consistency in determinations. In this case, perspective on the degree of confidence in the supporting WOE for the hypothesized mutagenic MOA for carbon tetrachloride (Table 5b) can be gained through comparison with the nature and extent of data available for the stronger database for TCP (Table 6).

Discussion

Comparative aspects of WOE analyses are illustrated here as a basis to contribute to transparency and consistency in delineating confidence/uncertainty in MOA/HR analysis based on the BH considerations. As noted by Guyton *et al.* (2008), Hill's (1965) considerations were not developed originally for evaluation of experimental/mechanistic data, though their utility for application in modified form to assess WOE in MOA analysis has been repeatedly though inconsistently tested. Based on increasing experience internationally in MOA/HR analysis (see, for example, Boobis *et al.*, 2006, 2008, Meek *et al.*, 2014), evolved B/H considerations are proposed here and clarified through delineation of the specific aspects addressed by each as framed by a series of questions. Definitions for these considerations have been additionally simplified and tailored to application in MOA analysis. The evolved B/H considerations were also rank ordered to reflect their relative contribution to WOE determinations and their utility exemplified in a comparative WOE approach.

The evolved B/H considerations build on previously published iterations and reflect experience in the application of MOA analysis. Several terms were clarified to facilitate assimilation of relevant chemical specific and biological data (i.e., "specificity" is now termed "essentiality of key events," "biological plausibility and coherence" is now termed "biological concordance" and concordance of empirical observations among key events delineated). In addition, considerations with limited relevance for evaluating MOA data (i.e., "strength," "coherence" and "experiment") were eliminated while other considerations (i.e., "analogy" and "incidence concordance") were added based on evolving experience with larger numbers of chemicals. It is hoped this evolved terminology, which reflects more common understanding within the broader risk assessment (versus epidemiological) community, will additionally contribute to consistency of use in MOA analysis. Finally, considerations were redefined as a basis to promote consistency and utility. For example, in publications of the IPCS MOA/HR framework (Boobis *et al.*, 2006, 2008; Sonich-Mullin *et al.*, 2001), consistency is defined as repeatability of key events in different studies; while in the USEPA cancer guidelines, consistency refers to the pattern of elevated risk observed across several independent studies (USEPA, 2005b). Neither definition accurately reflects the use of consistency in evaluating the WOE for hypothesized MOA(s). The former simply assesses reproducibility of results and, as such, may only contribute to the level of confidence in the occurrence of one key event. The latter definition is more appropriate to the assessment of the reproducibility of results in epidemiological and not mechanistic data sets. Consistency in the context of the MOA/HR framework more appropriately relates to evaluation of the WOE supporting interdependence of the key and end (adverse) events. Therefore, consistency was redefined here to reflect support of the pattern of effects across species/strains/organs and test systems for the hypothesized MOA. For example, if metabolism is a hypothesized key event in a carcinogenic MOA, the pattern of species-, strain- and sex-related variations in tumor response is compared to that expected based on known differences in metabolic profiles in the test systems. As such, it is not as important to assess if the occurrence of tumors is reproducible across studies, but rather, if the presence or absence of tumors in various species and strains is consistent with the hypothesized MOA.

Comparative WOE analysis is illustrated as a means of increasing understanding of the nature of transparency that is essential when evaluating confidence in the supporting WOE for hypothesized (potentially competing) MOAs. In doing so, it also provides a basis for increasing consistency in evaluation. Presentation of an overview of the data in a comparative manner (i.e., as supporting, inconsistent and missing) based on templates that cue evaluators concerning critical aspects provides concise insight into the extent of available data and relevant patterns in the existing database, which support various levels of confidence in considered options. In addition, this presentation concisely communicates areas of uncertainty (inconsistent data column and blank cells) and highlights areas of greatest impact for future research (missing data column). Ideally, further transparency on

the impact of this information (i.e. supporting, inconsistent and missing data) on the MOA conclusions would be provided in a detailed, supplemental discussion.

Synthesis of a collective data set to evaluate WOE for a hypothesized MOA is complex and challenging, requiring multidisciplinary input from both the research and risk assessment communities. This analysis is dependent upon transparent and consistent evaluation of the extent and nature of both chemical-specific and biological data versus supposition about possibilities for which there is essentially no experimental support. Characterization of the evolved B/H considerations is anticipated to contribute to more robust and transparent analyses, as a basis also to discourage, without clear rationale, the discounting of well-supported options based on the emphasis of outlying data of lesser quality.

This manuscript extends MOA/HR assessment through evolution of the B/H considerations and illustration of a comparative WOE analysis. Ultimately, it is anticipated that the additionally articulated and comparative aspects, which build on considerable recent experience in MOA analysis, will contribute to increasing transparency, consistency and methodological rigor in separating aspects of science judgment (i.e., weighting of options based on transparent consideration of available scientific support) from those of public policy in regulatory risk assessment (the latter of which sometimes involves embedded conservatism, to increase public health protection).

Acknowledgements

This work builds on previous products of a significant number of contributing authors and collaborators who have participated in the development and application of mode of action/species concordance analysis and associated training materials. Contributors include those outlined in Meek *et al.* (2003, 2014), Seed *et al.* (2005) and Boobis *et al.* (2006, 2008).

Conflict of Interest

Several of the authors (C.M.P., A.N.B., C.M.N. and R.J.L) are employed by a subsidiary of Exxon Mobil, who produces materials evaluated by the US EPA. Methodological aspects based on case studies considered here do not relate to specific evaluations of relevance to Exxon Mobil.

The work reported in the paper was conducted during the normal course of research/training for the University of Ottawa (MEM). It evolved from discussions during and following provision of training, for which funding for development of materials, travel and lodging was provided.

References

Boobis AR, Cohen SM, Dellarco V, McGregor D, Meek ME, Vickers C, Willcocks D, Farland W. 2006. IPCS framework for analyzing the relevance of a cancer mode of action for humans. *Crit. Rev. Toxicol.* **36**: 781–792.

Boobis AR, Doe JE, Heinrich-Hirsche B, Meek ME, Munn S, Ruchirawat M, Schlatter J, Seed J, Vickers C. 2008. IPCS framework for analyzing the relevance of a non-cancer mode of action for humans. *Crit. Rev. Toxicol.* **38**: 87–96.

Borgert CJ, Mihairch EM, Ortego LS, Bentley KS, Holmes CM, Levine SL, Becker RA. 2011. Hypothesis-driven weight of evidence framework for evaluating data within the US EPA's Endocrine Disruptor Screening Program. *Regul. Toxicol. Pharmacol.* **61**(2): 185–191.

Carmichael N, Bausen M, Boobis AR, Cohen SM, Embry M, Fruijtier-Polloth C, Greim H, Lewis RJ, Meek ME, Mellor H, Vickers C, Doe F. 2011. Using mode of action information to improve regulatory decision-making: an ECETOC/ILSI RF/HESI workshop overview. *Crit. Rev. Toxicol.* **41**: 175–186.

Dellarco VL, Baetcke K. 2005. A risk assessment perspective: application of mode of action and human relevance frameworks to the analysis of rodent tumor data. *Toxicol. Sci.* **86**: 1–3.

Guyton KZ, Barone S, Brown RC, Euling SY, Jinot J, Makris S. 2008. Mode of action frameworks: a critical analysis. *J. Toxicol. Environ. Health B* **11**: 16–31.

Hill AB. 1965. The environment and disease: An association of causation. *Proc. R. Soc. Med.* **58**: 295–300.

Manibusan MK, Odin M, Eastmond DA. 2007. Postulated carbon tetrachloride mode of action: a review. *J. Environ. Sci. Health.* **25**(3): 189–209.

Meek ME. 2008. Recent developments in frameworks to consider human relevance of hypothesized modes of action for tumours in animals. *Environ. Mol. Mutagen.* **49**(2): 110–116.

Meek ME, Klaunig J. 2010. Interpreting available data and identifying critical data gaps for benzene risk assessment. *Chem. Biol. Interact.* **184**, 279–285.

Meek ME, Bucher JR, Cohen SM, Dellarco V, Hill RN, Lehman-McKeeman LD, Longfellow DG, Pastoor T, Seed J, Patton DD. 2003. A framework for human relevance analysis of information on carcinogenic modes of action. *Crit. Rev. Toxicol.* **2003**(33): 591–653.

Meek ME, Berry C, Boobis AR, Cohen SM, Hartley M, Munn S, Olin SS, Schlatter J, Vickers C. 2008. Mode of action frameworks: a critical analysis. *J. Toxicol. Environ. Health B* **11**: 681–685.

Meek ME, Boobis A, Cote I, Dellarco V, Fotakis G, Munn S, Seed J, Vickers C. 2014. New developments in the evolution and application of the WHO/IPCS framework on mode of action/species concordance analysis. *J. Appl. Toxicol.* **34**: 1–18.

NTP. 1993. Toxicology and carcinogenesis studies of 1,2,3-trichloropropane (CAS No. 96-18-4) in F344/N/N rats and B6C3F1 mice (gavage studies). Public Health Service, U.S. Department of Health and Human Services; NTP TR 384. NIH Publication No. 94–2839.National Institute of Enviornmental Health Sciences: Research Triangle Park, NC.

SAB. 1999. Minutes of the Chloroform Risk Assessment Review Subcommittee Meeting, October 27–28, 1999.

SAB. 2007. Subject: Advisory on EPA's assessments of carcinogenic effects of organic and inorganic arsenic: a report of the US EPA Science Advisory Board. Available from: http://www.yosemite.epa.gov/sab/sabproduct.nsf/EPA-SAB-07-008.pdf.

SAP. 2000. Atrazine: Hazard and dose–response assessment and characterization, SAP Report No. 2000–05. FIFRA Scientific Advisory Panel Meeting, June 27–29, 2000. Available from: http://www.epa.gov/scipoly/sap/meetings/2000/june27/finalatrazine.pdf.

Seed J, Carney E, Corley ER, Crofton K, DeSesso J, Foster P, Kavlock R, Kimmel G, Klaunig J, Meelk E, Preston J, Slikker W, Tabacova S, Williams G. 2005. Overview: Using mode of action and life stage information to evaluate the human relevance of animal toxicity data. *Crit. Rev. Toxicol.* **35**: 663–672.

Sonich-Mullin C, Fielder R, Wiltse J, Baetcke K, Dempsey J, Fenner-Crisp P, Grant D, Hartley M, Knaap A, Kroese D, Mangelsdorf I, Meek ME, Rice J, Younes M. 2001. IPCS conceptual framework for evaluating a mode of action for chemical carcinogenesis. *Regul. Toxicol. Pharmacol.* **34**: 146–152.

USEPA. 2005a. Science Issue Paper: Mode of carcinogenic action for cacodylic acid(dimethylarsinic acid, DMAV) and recommendations for dose response extrapolation July 26, 2005, Health Effects Division, Office of Pestcide Programs, US Environmental Protection Agency. Available from: http://www.epa.gov/oppsrrd1/reregistration/cacodylic_acid/dma_moa.pdf.

USEPA. 2005b. Guidelines for Carcinogen Risk Assessment, EPA/630/P-03/001F. Risk Assessment Forum, US Environmental Protection Agency, Washington DC, March 2005.

USEPA. 2009. Toxicological Review of 1,2,3-Trichloropropane (CAS No. 96-18-4) in Support of Summary Information on the Integrated Risk Information System (IRIS), EPA 635/R-08/010F, US Environmental Protection Agency, Washington DC, September 2009.

USEPA. 2010. Toxicological Review of Carbon Tetrachloride (CAS No. 56-23-5) in Support of Summary Information on the Integrated Risk Information System (IRIS), EPA 635/R-08/005F, US Environmental Protection Agency, Washington DC, March 2010.